



# The Commonwealth as Agent: Group Action, the Common Good, and the General Will

## Citation

Schofield, Paul C. 2013. The Commonwealth as Agent: Group Action, the Common Good, and the General Will. Doctoral dissertation, Harvard University.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:11051217>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The Commonwealth as Agent:  
*Group Action, the Common Good, and the General Will*

A dissertation presented  
by

Paul Schofield

To  
The Department of Philosophy

in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy  
in the subject of  
Philosophy

Harvard University  
Cambridge, Massachusetts

October 2012



**The Commonwealth as Agent:  
Group Action, the Common Good, and the General Will**

**Abstract**

In this dissertation, I argue for a Rousseauvian vision of an ideal society: one in which the people constitute a group agent, unified under a collective will, willing action that constitutes the common good. Most have tended to believe that the contrasts between an individual agent and an entire people are stark, and so accounts of the commonwealth that appeal to group agency at all usually emphasize the differences between them. I will argue, however, that members of a society collectively constitute an agent that resembles an individual agent more closely than is normally supposed. Specifically, I will argue that a society is under normative pressure to engage in projects and activities that are good or worthwhile, that it may permissibly impose burdens on some members as it attempts to realize the common good, and that when properly constituted it possesses a collective will that has authority over its members.

I begin in Chapter 1 by giving an account of what it means for a group to engage in action. Then, in Chapter 2, I argue that the people of a region naturally constitute a rational group agent, and that state institutions that function properly help to facilitate collective action that is generally worthwhile or good. In Chapter 3, I argue that an individual has duties to herself similar to those that a commonwealth has to its members. For this reason, I suggest that investigating individual agency has the potential to shed light on what the commonwealth may or may not permissibly do. In Chapter 4, I draw a parallel between an individual agent pursuing her own good while trying to avoid wronging herself, and an entire commonwealth pursuing its overall good while trying to avoid wronging its members. There I conclude that by realizing the common good, the commonwealth compensates members who are burdened in its pursuit. In Chapter 5, I consider how a commonwealth, understood as a group agent, can choose its action,

arguing that democratic institutions possess authority over the commonwealth, and thus constitute the society's General Will. Finally in Chapter 6, I consider whether philosophers in the liberal-republican tradition have given sufficient reason for us to reject the Rousseauvian view that I argue for in the first five chapters.

## Table of Contents

<i>Acknowledgments</i> .....	vi
Introduction.....	1
1. Commonwealths as Agents? .....	1
2. Outline of the Project: Group Action, the Common Good, the General Will.....	5
Chapter 1: Group Action .....	8
1. Introduction.....	8
2. Bratman's Approach and an Alternative.....	9
3. The Circularity Problem .....	17
4. Group Intentions.....	22
5. Toward an Account of Group Action.....	30
Chapter 2: Society as a Natural Agent .....	37
1. Introduction.....	37
2. Natural Agency.....	38
3. The Difference Rationality Makes.....	46
4. The Relationship Between the Individual and the Group .....	50
5. Communitarianism .....	55
6. Society Can Determine What Members Do .....	58
7. The Commonwealth as a Natural Agent.....	68
Chapter 3: Duties to the Self.....	71
1. Introduction.....	71
2. Morality's Second-Personal, Bipolar Character.....	73
3. The Possibility of Relating to Oneself Second-Personally .....	77
4. Duties to the Self .....	83
5. The Significance of Moral Reasons .....	87
Chapter 4: The Common Good .....	89
1. Introduction.....	89
2. Compensation and Justification .....	90
3. Compensation and the Unity of Action.....	93
4. The Common Good and Compensation.....	99
5. Worthwhile Group Action as a Good .....	101
6. The "Intertwining" of Interests .....	107
7. Answering Some Worries about the Very Idea of a Common Good .....	109
8. Conclusion.....	113
Chapter 5: The General Will .....	114
1. Introduction.....	114
2. The Service Conception.....	114
3. The Transparency Conception of Authority.....	118
4. Establishing Authority .....	124
5. Political Judgment and Collective Activity .....	126
6. Taking One's Own Political Judgment to Be Decisive .....	129
7. Equal Voice and Democracy.....	137
8. A Possible Dilemma .....	141
9. Conclusion.....	143
Chapter 6: Liberalism, Republicanism, and Choice .....	144
1. Introduction.....	144
2. Ripstein's Kantian View .....	146
3. Rawls's Political Liberalism .....	153
4. Conclusion.....	157

## Acknowledgments

I owe thanks to many people for their help and encouragement with this project. I would like to express gratitude to my committee for their guidance throughout the process. Chris Korsgaard, Dick Moran, and Tim Scanlon were as wonderful a group of advisors as one could hope for. I'd also like to thank Selim Berker, Matt Boyle, Ned Hall, Doug Lavin, Tommie Shelby, and Susanna Siegel—all members of a terrific department that helped make me a better philosopher. During my years at Harvard, I benefitted greatly from the excellent graduate student community that we have here. For their friendship and for all the philosophical conversation, I owe a special debt of gratitude to Ryan Doerfler, Patricio Fernandez, Nathan Foell, Tom Fehse, Jeremy Fix, Chris Furlong, Micha Glaeser, Paul Julian, Michael Kessler, Dave Langlois, Craig Nishimoto, Kristi Olson, Nataliya Palatnik, Bill Porter, and Jiewuh Song. I would also like to thank Jeffrey Schneibel, Carla Bagnoli, and Luca Ferrero for encouraging me early on.

I am very grateful to my parents, who sacrificed for the sake of my education, and who supported me even when they had no idea what I was doing. Penelope and Truman arrived late in the dissertation writing process. Both provided a happy distraction during down time as I completed writing. Finally, and most of all, I would like to thank my wife Maureen for all her love and support throughout this process. She has shown tremendous patience. I couldn't have done this without her.

*For Maureen*



# Introduction

## 1. Commonwealths as Agents?

In this dissertation, I argue that a commonwealth—that is, a political community—is an agent. I show that a group of people living in a geographical area, who interact and come into causal contact with one another, constitute a natural agent. That is to say, such a group *just is* an agent, regardless of any decision anyone has ever made, and regardless of any explicit attempts on the part of the people to organize themselves into a unified whole. Once they do organize themselves, however—once they create state institutions—the people becomes a fully-fledged commonwealth, that not only is an agent, but that actively exercises its agential capacities.

That a people can act together as a group is an idea that has enjoyed a storied career in the history of political thought. Advocates for democracy and republicanism have frequently suggested that a legitimate, well-constituted government is one that enables a people to collectively rule itself, making decisions and carrying them out together as a whole. The concept of group agency has been integral to the political philosophies of thinkers as different as Plato, Hobbes, Locke, Rousseau, Kant, Hegel, and probably Rawls.<sup>1</sup> And even the American Constitution employs the concept of collective agency, as its ratification and amendment are, supposedly, carried out by *the people* through procedures that enable such acts to be attributed *to them*.<sup>2</sup> An account that makes use of this idea—the idea of a people being an agent—must rely, at least implicitly, on views about what it means to say that a people is an agent, and about what the normative implications of its being an agent are. Most have tended to believe that in these regards, the contrasts between individual agents and entire commonwealths are stark, and so

---

<sup>1</sup> Joshua Cohen reports that Rawls once remarked that *A Theory of Justice* was an attempt to “spell out the content” of Rousseau’s general will, suggesting that Rawls believed that the principles of justice were supposed to be willed and implemented collectively. (See Joshua Cohen, *Rousseau: A Free Community of Equals* (Oxford: Oxford University Press, 2010), p. 2.)

<sup>2</sup> See Bruce Ackerman, *We the People, Volume 1: Foundations* (Cambridge, Mass.: Harvard University Press, 1991), p. 6.

accounts of the commonwealth that appeal to group agency often emphasize the differences between an individual and a people. However, I will argue that members of a society collectively constitute an agent that resembles an individual agent more closely than is normally supposed.

To elaborate a bit on one of the most significant similarities that I will highlight, consider how individual agents are usually thought to be under normative pressure to live out lives of worthwhile activity—to, quite generally, find things that are worth doing, and then do them. A political body, by contrast, is usually thought to be more limited in its aims. Most groups and corporate bodies, after all, are not under rational pressure to do worthwhile things generally. For instance, the board of directors at a major company can admit that feeding the hungry, putting on a ballet, and preserving the wetlands are all things that are worth doing, while at the same time acknowledging that the company itself is under no normative pressure to undertake these projects. The purpose of a business, one might think, is simply to make a profit, and not to do good things generally. It is often said that a commonwealth's proper aims are similarly limited. Its proper pursuits are said to include only things like protecting its members' rights, maintaining a civil society that is free, providing essential services, and mediating disputes. The mere fact that some endeavor is good or worthwhile does not, in itself, constitute even a *prima facie* reason for the commonwealth to involve itself in it. It is this view that I wish to reject. I will argue, instead, that the fact that a particular collective action is good or worth doing does indeed count as a *prima facie* reason for the commonwealth to do it—I argue that human societies are under normative pressure generally to find worthwhile things to do, and that it is part of the role of state institutions to help facilitate such pursuits.

In addition to this purported difference, it is also typically assumed that individual agents and commonwealths differ in that the former may permissibly act in many ways that the latter may not. Most liberals, for instance, believe that individual agents may impose burdens upon themselves at particular moments in order to advance their overall good. They may, for instance, undergo painful medical procedures, take on onerous financial obligations, and

sacrifice spending time with their family, in order to do things that will eventually make their lives go better overall. This is, most would agree, simply part of what living a life consists in. Most deny, however, that entire commonwealths may permissibly seek to advance their overall good by imposing burdens on individual members—or at least, most would deny that doing so is as morally innocuous as an individual’s imposing a burden upon herself.<sup>3</sup> It would certainly be dubious, after all, to suggest that there is no moral impediment at all to substantially increasing someone’s taxes, or disposing of waste in someone’s neighborhood, in order to help advance the overall social good. These burdens need to be justified to the individual who suffers them, and it isn’t self-evident that advancing the overall social good is sufficient to render such burdens justifiable. This is because while a burdened individual can be compensated through contributions to her overall good, an individual, it is often said, *cannot* be compensated for burdens that she bears through a contribution to society’s overall good. The taxed individual, for instance, isn’t *herself* compensated through a contribution to society’s good. Rather, something is taken from her and is given to someone else. Against this, I will argue that compensating an individual through a contribution to society’s overall good is, in fact, possible, and that for this reason a commonwealth is permitted to pursue its overall good in much the same way that an individual is.

To mention one more way that an individual and a commonwealth are thought to differ, consider that most liberals and republicans believe that an individual agent has a *prima facie* moral claim to decide for herself which activities she will engage in, and which ends she will pursue—and also a claim to not be drawn into the ends of others. Philosophers differ on the details. But both the liberal and republican traditions have tended to deny that entire societies have a general moral permission to pursue collective projects and ends of their choosing, as

---

<sup>3</sup> Rawls famously believes that individuals are permitted to impose burdens upon themselves in order to pursue their overall good, but criticizes utilitarians for thinking that entire societies are permitted to act similarly. (See his *A Theory of Justice, Revised Edition* (Cambridge, Mass.: Harvard University Press, 1999), p. 24.)

doing so would require enlisting individuals, who have a right to decide for themselves which projects to engage in. The purpose of state institutions, it is often thought, is merely to create an environment within which rational individuals may live lives of their own choosing, while clashing with one another as little as possible. Choosing collective ends, and facilitating the collective pursuit of those ends, is not part of the state's role, and a commonwealth oversteps its bounds when it uses state institutions to do so.<sup>4</sup> Against this, I will argue that a whole society, when it is democratically constituted, possesses the authority to enlist individuals to carry out collective plans, ends, and projects that it deems to be good or worthwhile, and that coercing an individual in order to enlist her in society's collective endeavors is often permissible.

On the whole, this dissertation can aptly be understood as an extended argument for a Rousseauvian vision of an ideal society: one in which the people constitute a group agent, unified under a collective will, willing action that constitutes the common good.<sup>5</sup> This vision differs in significant ways from the ones that liberals and republicans tend to favor. My goal, however, is not to offer a wholesale critique of liberal or republican thought, as some communitarians have done.<sup>6</sup> Rather, I take myself to share with liberals and republicans many fundamental commitments, such as their belief in self-rule and their belief that government policy must be justifiable to each and every individual. Nevertheless, I want to show that these commitments are consistent with—and indeed can sometimes be enlisted in support of—a vision of the commonwealth very different from the that one most liberals and republicans have tended

---

<sup>4</sup> I do not mean to quibble here about whether willing laws designed to make individuals free to pursue their own ends counts as “choosing an end.” My point is just that most liberals and republicans do not think that the commonwealth is in the position of having to identify and choose valuable projects and ends to pursue generally.

<sup>5</sup> Rousseau's ideal commonwealth is a collective body constituted by all of society's members, each of whom possesses a general will with which she wills the common good. My debt to Rousseau is significant, but there are deep differences between the view I argue for in this dissertation and Rousseau's. To name two prominent differences, the notion of a social contract plays no role at all on my account, and freedom plays a less central role on my view than it does on Rousseau's.

<sup>6</sup> Michael Sandel, for instance, urges his version of communitarianism not as an extension of liberal principles, but as an alternative to them. (See his *Democracy's Discontent: America in Search of a Public Philosophy* (Cambridge, Mass.: Harvard University Press, 1996).)

to advocate. I hope, then, to capture much of what is appealing about liberalism and republicanism. But at the same time I hope to put forth a view that avoids the deficiencies that Marx thought he saw in these traditions when he famously accused them of regarding each person, “as an isolated monad, withdrawn into himself.”<sup>7</sup>

## 2. Outline of the Project: Group Action, the Common Good, the General Will

As the title of the dissertation suggests, three concepts will play a particularly important role in my account: *group action*, the *common good*, and the *General Will*. By explicating each, I hope to gain insight into the nature of an entire people’s agency, and ultimately propose answers to some perennial philosophical questions, such as what the purposes of state institutions are, how policies that burden some individuals in order to benefit others can be justified, and why citizens should obey the law.

*Group action* has recently been the subject of much philosophical discussion. I begin in Chapter 1 by clarifying the problem that I see with some prominent accounts and arguing for an alternative view. Then in Chapter 2, with my own account of group action having been spelled out, I argue for my central claims, which are that the people of a region naturally constitute a rational group agent, and that properly functioning state institutions help to facilitate collective action that is generally worthwhile or good.

Even if I successfully establish that the commonwealth is a rational agent, it will be objected that the commonwealth is severely limited in what it may permissibly do, because engaging in action that is worthwhile overall might require burdening some individuals, or enlisting them against their wills in a collective endeavor with which they disagree. For these reasons, it is commonly supposed that political bodies are morally constrained in ways that individual agents are not. So in Chapter 3, I demonstrate that for an individual to impose

---

<sup>7</sup> Karl Marx, “On the Jewish Question,” in *The Marx-Engels Reader (Second Edition)*, ed. Robert Tucker (New York: Norton, 1978), pp. 26-46 (quotation from p. 43).

burdens upon herself, and to undermine her own autonomy, can also be morally problematic. In this way, the individual and the commonwealth are, I argue, more alike than is typically acknowledged. And if despite having moral obligations to herself the individual may permissibly balance burdens and benefits in order to pursue her overall good, or if she may permissibly set her own ends, then it raises the possibility that the commonwealth may do these things as well.

In Chapter 4, I draw a parallel between an individual agent pursuing her own good while trying to avoid wronging herself, and an entire commonwealth pursuing its overall good while trying to avoid wronging its members. I begin by arguing that imposing burdens upon oneself is sometimes morally permissible because compensation occurs when one's overall good is advanced. I then proceed to make a similar argument about the commonwealth. I argue that a commonwealth's doing worthwhile things constitutes the *common good*, which is to say, a good or interest of each individual member. This argument will allow me to claim that by acting in ways that are good, the commonwealth simultaneously promotes the good of each individual member, thus compensating members for burdens that they bear.

In Chapter 5, I consider how a commonwealth, understood as a group agent, can choose its action. There I argue that democratic institutions possess authority over the commonwealth, and thus constitute the society's *General Will*. I defend the claim that citizens should obey democratically enacted law by arguing that democratic institutions simply *are* the General Will—their operation is a people's way of deciding for themselves what to do. Just as a rational individual often does what *her* will settles upon, I argue that rational citizens do as the General Will decides when it makes decisions about how citizens will act within the public sphere—I argue that members obey democratically enacted law insofar as they are rational. In this sense, the democratically constituted commonwealth has authority, and *authors* the actions carried out by the whole.

Finally in Chapter 6, I conclude by considering a set of views that hold that state institutions exist merely to create an environment within which individuals can carry out

individual activities and projects of their own choosing, rather than to help facilitate collective action or facilitate the realization of collective ends. This set of views constitutes a rival to the general view that I develop over the course of this dissertation. I conclude by giving a quite general reason for calling this rival conception into question.

## Chapter 1: Group Action

### 1. Introduction

The thesis of this dissertation is that the commonwealth is an agent that engages in intentional action. What kinds of things might a commonwealth do? Some promising candidates include going to war, putting in place and sustaining a socialist economy, realizing and sustaining a property-owning democracy, building and maintaining the Eisenhower Interstate Highway System, recognizing the institution of marriage, amending the American Constitution, and realizing the common good. These rather grand and large-scale actions are species of the genus “group action,” which includes more modest items, such as a group’s shopping for clothes and a group’s travelling to New York. If we hope to better understand what happens when an entire people goes to war, realizes the common good, amends the Constitution, and so on, then we must first inquire into the metaphysics of the genus—inquire, that is, into the metaphysics of group action.

English permits us to say “we travelled to New York” even when you and I were unaware of the other’s existence at the time of the trip. Here, “we travelled” means merely that both you travelled and I travelled. English similarly permits us to say “we painted the house” when you and I each painted half, unbeknownst or indifferent to the other. Here, “we painted it” means merely that you and I each painted part. When philosophers talk about “group action,” however, they are generally interested in a different phenomenon: action that is done *cooperatively* or *together*.<sup>1</sup> That will be my topic in this chapter.

One way to give an account of group action is to identify attitudes that must accompany a bodily movement—attitudes that can be defined without appealing to the concept of group

---

<sup>1</sup> Nearly everyone writing on this topic begins by distinguishing between individual actions in aggregate on the one hand, and cooperative action (or action done together) on the other. For clear illustrations of this distinction, see John Searle, “Collective Intentions and Actions,” in *Intentions in Communication*, eds. P. Cohen, J. Morgan, and M. Pollack (Cambridge, Mass.: MIT Press, 1990), pp. 401-416 (esp. 404-405); Margaret Gilbert, *On Social Facts* (Princeton: Princeton University Press, 1989), pp. 154-157.



action itself—in order for the movement to be a group action. Another way is to identify mental states or states of affairs that must cause a bodily movement in order for it to be a group action. In this chapter, I will argue against these approaches. While I cannot prove that no such approaches could ever succeed, I will demonstrate that they run into significant difficulties. I will then suggest an alternative that I believe fares better.

## 2. Bratman's Approach and an Alternative

“Group action” and “group intention” are often treated as related, but distinct philosophical topics.<sup>2</sup> The former pertains to deeds done by a group, whereas the latter pertains to mental states of an individual, inter-related mental states of several individuals, or other states of affairs that are responsible for bringing about such deeds. I will begin by discussing group action, and move on to group intention later.

One approach in the group action theory literature is to begin with the more general concept of an action, and then identify features that any action must possess in order to be a *group* action. Michael Bratman, for instance, starts with action-types such as *travelling to New York* and *painting the house*, and proceeds to isolate features distinctive of those instances in which such actions are done *cooperatively*. To spell this out a bit more, he first insists that genuine group action—genuine instances of, say, *travelling to New York*—must be distinguished

---

<sup>2</sup> For instance, both Margaret Gilbert and Michael Bratman begin by addressing the topic of group action (or “shared cooperative activity” or “acting together”), and only later turn to analyzing group intention. This division within the field mirrors a common division in the action theory literature more generally. It is common for action theorists to discuss intentions on the one hand, intentional action on the other, and then investigate the connection between the two. This is what we see in the group action literature as well. Early on Bratman published “Shared Cooperative Activity,” following it up later with “Shared Intention.” (Both essays are contained in his *Faces of Intention* (Cambridge, Mass.: Cambridge University Press, 1999).) In a recent paper, Gilbert says that while her early work focused on “acting together,” her later work addresses the topic of shared or group intentions. (See her “Shared Intentions and Personal Intentions,” *Philosophical Studies* 144 (2009), pp. 167-187 (quotation from p. 167).) This is not to suggest that Gilbert and Bratman believe that these topics are sharply distinct. I mention this simply to clarify why I will begin with group action, ignoring group intention until later.

from processes carried out by unsuspecting collections of persons.<sup>3</sup> To this end, he says that each participant must attempt to guide her behavior so as to be responsive to the others' behavior, while believing that the others will do the same.<sup>4</sup> But Bratman notes that this isn't sufficient to distinguish group actions from other actions. Soldiers on opposite sides of a conflict will also attempt to guide their behavior so as to be responsive to the others' behavior, while believing that those on the other side will do the same.<sup>5</sup> And certainly we don't want to call *that* "group action," in the sense of shared cooperative activity. So, recognizing that responsive behavior—or attempted responsive behavior—is a feature common to the warring soldiers and the cooperating travelers, Bratman looks for an additional feature that would exclude the warring soldiers. The feature he settles on is a "commitment to mutual support."<sup>6</sup> With this, the inquiry takes a particular shape: Bratman proceeds by identifying further features, the possession of which renders an action a *group* (or shared cooperative) one. Ignoring much of the nuance in his discussion of each necessary condition, we can state Bratman's view in the following way: Group actions are those actions in which there is (1) mutual responsiveness, (2) commitment to joint activity,<sup>7</sup> and (3) commitment to mutual support.<sup>8</sup>

Certainly it is helpful and informative to analyze *some* concepts this way. Unfamiliar with the notion of a dry town, I ask you to explain. Knowing that I grasp the concept of a town, you identify the feature that distinguishes dry towns from the rest, saying, "They are towns with a prohibition on alcohol." Understanding is thus achieved. Using Anton Ford's helpful

---

<sup>3</sup> Bratman prefers to talk of "shared cooperative activity," though I will simply talk of "group action" throughout.

<sup>4</sup> Bratman, "Shared Cooperative Activity," p. 94.

<sup>5</sup> Ibid., p. 95.

<sup>6</sup> Ibid.

<sup>7</sup> Bratman thinks "joint activity" can be characterized in a cooperatively neutral way, so he does not think that this condition renders his account circular.

<sup>8</sup> Ibid., pp. 94-95.

terminology, we can say that “dry town” is an *accidental species* of the genus “town.” When a species is accidental, “the contrast between [it] and its genus is mediated by an independent quality, or ‘accident.’”<sup>9</sup> In the present case, the mediating accident is alcohol prohibition—a feature that might also be had by a county, a campus, or a nation. If group action is an accidental species of the genus “action,” with the contrast between it and its genus mediated by accidental features (such as the presence and causal efficacy of various mental states and attitudes), then the appeal of analyzing group action in the way that Bratman does is obvious.

Not every species-genus relationship is accidental, however. An example of a species-genus relationship that is arguably not mediated by an accident is the relationship between a particular color and the genus “color.” The concept “green,” one might think, is not defined as the subset of colors that have an accidental feature, the possession of which sets them apart from the other colors. Green differs from other colors, we might say, *in respect of color itself*. Putting the point in different terminology, it is sometimes said that color is “determinable” and that a particular color is “determinate,” where the relationship of a determinable to a determinate is like the relationship of a genus to a non-accidental species that belongs to it. (Some believe that the determinable-determinate relationship should be sharply distinguished from the species-genus relationship, though I will not do so here.)<sup>10</sup>

Though more controversial, some contend that animal species are non-accidental species of the genus “animal.” Ford writes:

It is true that a horse differs from an octopus in various respects, and that  
in many of these it differs from an iceberg: e.g., in respect of its weight,

---

<sup>9</sup> Anton Ford, “Action and Generality,” in *Essays on Anscombe’s Intention*, eds. Anton Ford, Jennifer Hornsby, and Frederick Stoutland (Cambridge, Mass.: Harvard University Press, 2011), pp. 76-104 (quotation from p. 83).

<sup>10</sup> W.E. Johnson, who introduced this terminology, treated the relationship between determinables and determinates as a kind of species-genus relationship in his *Logic, Pt. I* (Cambridge, UK: Cambridge University Press, 1921). For an example of someone who believes that the two should be kept distinct, see John Searle, “On Determinables and Resemblance, II,” *Proceedings of the Aristotelian Society* (Supplementary Volume), XXXIII (1959), pp. 141-158. This dispute matters very little for my purposes.

smell, texture, volume, temperature, and flammability. But a horse also differs from an octopus in ways that it could only differ from another animal: e.g., in respect of its organs and members, and in respect of its manner of nourishing itself, and of reproducing, and of moving itself from place to place. The latter are differences internal to the genus. Just as what is red and what is blue differ in respect of color. . . so, too, what is a horse and what is an octopus differ in respect of animality.<sup>11</sup>

It is, however, more tempting to treat “octopus” as an accidental species of the genus “animal” than it is to treat “green” as an accidental species of the genus “color.” One might reasonably attempt to define “octopus” by identifying accidental features that differentiate it from other animals—beginning by pointing out that the octopus has eight legs, and then proceeding to pick out differentiating features until it is fully distinguished from spiders and other octopeds. However, those who object to treating the octopus as an accidental species deny that the octopus’s legs can be appealed to without appealing to the concept “octopus.” The octopus, after all, does not simply have eight legs. It has eight *octopus* legs. Likewise, a horse has *horse* legs, and a spider has *spider* legs. That is to say, the parts that constitute an animal of a particular species are not accidents that can be defined or understood without appealing to the role that they play in constituting a member of the particular species itself.<sup>12</sup> They are essentially parts of the whole that they constitute.<sup>13</sup> Of course, we do have the general concept “leg.” But it isn’t enough, when defining “octopus,” to appeal simply to the general concept, as if *just any*

---

<sup>11</sup> Ford, “Action and Generality,” pp. 86-87.

<sup>12</sup> Ibid., pp. 88-90.

<sup>13</sup> Aristotle went so far as to claim that a hand severed from a human body is not a hand at all, “except homonymously.” He thinks the relationship between a human and her members, or a human and her organs, is one in which, “the whole is necessarily prior to the part.” (*Politics*, trans. B. Jowett, *The Complete Works of Aristotle, Volume Two*, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984), p. 1988 (1253<sup>a</sup> 15-30).)

appendages that met the definition of a leg would suffice. An octopus must possess legs appropriate to its species.<sup>14</sup>

When explicating a concept, the kind of account that it is proper to give will very much depend upon whether a species is accidental or non-accidental. If S is an accidental species of G, then we can give a proper, non-circular definition of S by specifying which accidents differentiate S from other members of G. Bratman clearly tries to give an account of this form, as the features that he thinks differentiate group actions from other actions are attitudes and mental states that can be defined without appealing to the concept of group action itself. So if it turns out that group action is a non-accidental species of the genus “action,” then Bratman’s account would be undermined.

So, I would like to raise the possibility of treating the genus “action” as *determinable*, roughly analogous to the way that “color” perhaps is, or the way that Ford believes “animal” is. Group action and individual action, on this view, would be *determinates*. So, just as red is perhaps distinguished from green by differences in respect of *color* itself (and not by the presence of some accident), and just as an octopus is perhaps distinguished from a horse by differences in respect of *animality* itself (and not by the presence of some accident), group action would be distinguished from individual action by differences in respect of *action*—though it will become clear what exactly this means only once I offer my positive account at the end of the chapter.<sup>15</sup> I suspect some readers will be skeptical of the viability of any view that treats

---

<sup>14</sup> Other common examples of genera whose species are accidental (that is, other examples of determinables) include shape and pitch. Michael Thompson has also recently argued that a “life-form” is not an accidental species of the genus “substance.” (See his, “The Representation of Life,” in *Life and Action: Elementary Structures of Practice and Practical Thought* (Cambridge, Mass.: Harvard University Press, 2008), pp. 25-84.)

<sup>15</sup> Caution is in order here. There is a substantial literature on the topic of determinables and determinates, and there exists disagreement about how to understand certain paradigm cases (color, shape, pitch, and so on). My remarks here are not intended to suggest that group action should be treated exactly as colors or as animals should. My intention is simply to gesture toward a general form that the relationship between the more general and the more particular sometimes takes, and to raise the possibility that the relationship between action and group action takes this form.

group action as a non-accidental species of the genus “action.” In this section, I will attempt to dispel some of that skepticism.

As with animal species, it is tempting to treat group action as an accidental species of its genus, because, as Bratman’s work shows, there are plausible candidates to mediate between “group action” and “action” (such as attitudes and mental states). The danger, though, is that some of the necessary conditions that he uses to give his analysis quietly appeal to the species itself—quietly appeal, that is, to the concept “group action.” For instance, something like “group action responsiveness” (and not merely “responsiveness”) might be the *real* necessary constituent of group action, just as the species-specific “octopus leg” (and not merely “leg”) is arguably the *real* necessary constituent of an octopus.<sup>16</sup>

My suggestion, then, is that what seem to be accidental features that could help reductively analyze group action might turn out not to be accidents at all, but essentially parts of the whole that they help constitute. Here I want to say a bit more to convince the reader that this is indeed a *legitimate* possibility, even if the reader is yet to be convinced that it is the *correct* one. To do so, I will draw some parallels with debates in individual action theory. This will require a slight detour.

Elizabeth Anscombe claims that the mark of a bodily movement’s being intentional is that, “a certain sense of the question ‘Why?’ is given application.”<sup>17</sup> If you ask why I am tapping my foot, and I sincerely answer, “I didn’t realize I was,” then the question is refused application, indicating that the tapping is not an intentional action. If I sincerely answer that I am keeping a beat, or sending a message in Morse Code, then I accept the applicability of the question,

---

<sup>16</sup> To agree with this, one needn’t deny that we have a general concept of “attempting to behave responsively,” which might be a feature common to the cooperative travelers, the warring soldiers, and my cat as she follows me around the house. One simply must deny that *just any* way of “attempting to behave responsively” will satisfy the necessary condition for group action, just as we denied that *just any* appendage that meets the definition of a leg can help constitute an octopus.

<sup>17</sup> Elizabeth Anscombe, *Intention, Second Edition* (Cambridge, Mass.: Harvard University Press, 2000), p. 9.

indicating that the tapping is indeed an intentional action. The “Why?” question not only helps differentiate intentional actions from mere bodily movements, but also helps differentiate one intentional action from another. Two people might each lightly stomp on my foot. When this happens, it may seem that two identical events occur. But when confronted with the “Why?” question, one foot-stomper sincerely answers: “I was getting your attention in order to flirt with you.” The other sincerely answers: “I was distracting you in order to pick your pocket.” While the internal constitutions of the movements are identical, the responses indicate that two distinct actions are performed and that two distinct events occur. One is a bit of a seduction. The other is a bit of a robbery. This marks a difference in the essence of what is being done when each individual stomps on my foot.

All agree that in stomping, the flirt and the thief do different things. But there are significant disagreements about how to differentiate the two. Many philosophers understand the intention to flirt to be a mental state that is causally responsible for action, and believe that we can differentiate a bit of flirting from a bit of larceny by appealing to the mental state’s intentional content. However, Anscombe—along with an increasing number of philosophers influenced by her—rejects this view. Of course, Anscombe agrees that the flirtation and the theft are done with different intentions. But she explicitly denies that an action or bodily movement is made intentional by, “a mere *extra* feature of [the event] whose description would otherwise be the same, only if we concentrate on small sections of action.”<sup>18</sup> Anscombe denies, for instance, that the flirt’s action can be distinguished from a mere episode of bodily movement by appealing to the fact that it is caused by a state of the mind that can be defined and understood without appealing to the concept “intentional action” itself. She similarly would deny that we can differentiate what the flirt does from what the thief does by appealing to the fact that their foot-stompings were caused by mental states containing different intentional content. What I wish to highlight here is simply Anscombe’s denial that an intention is a “mere extra feature” of an

---

<sup>18</sup> Ibid., p. 88.

intentional action. This marks an important divide among action theorists: those who treat intention as a “mere extra feature,” and those, like Anscombe, who do not.

I will say more about this in section 4, but for the moment let us return to group action. Just as we can imagine two seemingly identical bodily movements that nevertheless differ with respect to the intention with which they are done, we can imagine two seemingly identical events in which a number of people engage in intentional action that nevertheless differ in that one is a group action and the other is not. John Searle gives the following example:

Imagine that a group of people are sitting on the grass in various places in a park. Imagine that it suddenly starts to rain and they all get up and run to a common, centrally located, shelter. Each person has the intention expressed by the sentence “I am running to the shelter.” But for each person, we may suppose that his or her intention is entirely independent of the intentions and behavior of others. In this case, there is no collective behavior; there is just a sequence of individual acts that happen to converge on a common goal. Now imagine a case where a group of people in a park converge on a common point as a piece of collective behavior. Imagine that they are part of an outdoor ballet where the choreography calls for the entire *corps de ballet* to converge on a common point. We can even imagine that the external bodily movements are indistinguishable in the two cases; the people running for shelter make the same types of bodily movements as the ballet dancers. Externally observed the two cases are indistinguishable, but they are clearly different internally.<sup>19</sup>

Bratman’s reductive account would distinguish between the two cases that Searle lays out by appealing to the presence of various states of mind that render what happens in the latter case a group action. But I want to suggest that, as was true with the seemingly identical foot-

---

<sup>19</sup> Searle, “Collective Intentions and Actions,” p. 403.



stompings, there are actually two legitimate approaches one might take here. The alternative to Bratman's approach is to insist that an action's being a *group* action isn't a matter of its having some extra ingredient or ingredients. That is to say, one might deny that "groupishness" is a mere extra feature of a group action, just as Anscombe denies that an intention is a mere extra feature of an intentional action. Group action could be, we might say, irreducibly "groupish." Thus, we should expect there to be two camps here, just as in individual action theory.

Though I have only barely hinted at what the positive, more Anscombian account would look like, it is nevertheless significant that such an approach might exist, and that Bratman never seems to consider this possibility or argue against it. In the next section, I will argue that there is actually positive reason for rejecting Bratman's approach.

### 3. The Circularity Problem

Many theories of group action have been accused of being circular.<sup>20</sup> In this section, I will advance a version of this objection. I will argue that some action-types are "irreducibly cooperative," and that Bratman's theory is ill-equipped to give a non-circular analysis of them. On Bratman's view, a commitment to joint activity is a necessary condition for group action. If he is to avoid very tight circularity, Bratman must characterize what the agent is committed or willing to do in a way that does not itself appeal to the concept "group action." Bratman tries to avoid circularity by attempting to characterize joint acts in ways that are "cooperatively neutral." Two kinds of acts seem capable of being done "jointly" but "non-cooperatively." The first are goal-directed acts. If one paints a house, one pursues a state of affairs in which the house is painted. We can stipulate that painting the house "jointly" means that two individuals intentionally make contributions to this state of affairs' being realized. Since two individuals might contribute without knowing about the other's contribution—if, for instance, they are

---

<sup>20</sup> For a good discussion of the circularity objection, see Björn Petersson, "Collectivity and Circularity," *Journal of Philosophy* 104 (2007), pp. 138-156.

painting opposite sides of the house—it is possible to paint the house jointly, but not cooperatively. The second are acts that an individual can carry out by herself, for their own sakes. If an individual travels by train to New York simply because she takes pleasure in the ride, she travels for its own sake. We can stipulate that travelling is done “jointly” when two individuals travel, from start to finish, in the presence of one another. Here too it is possible for two people to travel to New York in one another’s presence without cooperating—as when strangers happen to purchase tickets for adjacent seats. If we consider only these two kinds of acts, we are likely to conclude that Bratman can avoid the circularity problem.

However, a difficulty arises for what I will call “irreducibly cooperative” acts. Take *dancing the tango*. *Dancing the tango* is not something done in pursuit of an end external to itself. There is no goal to which each participant makes a contribution, other than the goal of *our dancing the tango* itself. Furthermore, an individual cannot dance the tango by herself, so it is impossible to imagine agents doing it “jointly” but non-cooperatively—for instance, by dancing individually for its own sake in the presence of one another, without cooperating. To dance the tango at all is to engage with another in a cooperative endeavor. If this is right, it spells trouble for Bratman. “A commitment to dance the tango” appeals to the notion of a cooperative endeavor.

The existence of irreducibly cooperative acts also threatens to undermine another part of Bratman’s view. Bratman argues that a “commitment to mutual support” is necessary for group action. Once again, this condition is best suited to help analyze acts that are goal-directed or acts that could be performed by individuals for their own sakes. If my partner drops her paint brush, I support her by picking it up and handing it to her. “Support” is whatever contributes to the end’s being realized, and handing her the brush facilitates the house’s coming to have a new coat of paint. Or if my partner repeatedly drops her heavy suitcase in the train station, I support her by carrying it myself. “Support” is whatever contributes to our remaining in one another’s presence as we travel to New York. But we encounter difficulty once again when trying to

characterize what “support” amounts to for group actions such as *dancing the tango*. Here “support” can only be characterized as something done in the service of *dancing the tango* itself. Since *dancing the tango* cannot be characterized in a cooperatively neutral way, and since “support” in the present case can only be characterized by helping ourselves to the notion of *dancing the tango*, it seems that Bratman’s “commitment to mutual support” condition falls prey to the circularity objection as well. And, though I won’t do it here, we could, I think, raise further questions about whether the other features that Bratman appeals to in his analysis also fall prey to the circularity objection.

With acts like *dancing the tango* in view, we should now ask whether the other kinds of action that we mentioned earlier are sometimes best understood as constituted by acts that are irreducibly cooperative, rendering Bratman’s attempt to analyze them tightly circular as well. Imagine a case in which an overbearing father promises his son that they will paint a house together—seemingly a goal-directed activity. The father then repeatedly touches up his son’s work, guides his son’s hand as he tries to paint, and micromanages every aspect of the project. Both father and son make contributions to the goal’s being realized, both are aware of the other’s contributions, and the father supports the son if what we mean is simply that he “facilitates the son’s making a contribution to the goal.” But at the end of the day, with the house painted, we can imagine the son intelligibly complaining: “But I thought *we* were going to paint the house.” What can we make of this complaint? Perhaps the son took their “commitment to painting the house” not merely to be a commitment to making a contribution to the house’s having a new coat of paint, but to be a commitment like the one that the dancers must have when they commit to dancing the tango—perhaps he took it to be a commitment to painting, understood in a sense that is irreducibly cooperative. And if he did, we can plausibly imagine that part of his point is that not just any way that the father could facilitate the son’s making a contribution to the goal would count as facilitating *painting the house* in the sense that he has in mind. Perhaps the son’s complaint is that the overbearing father’s purported “support”—the

“support” that seems overbearing—actually undermines the cooperative endeavor, rather than helps constitute it. If the action is irreducibly cooperative, then it seems possible for there to be ways of pursuing the goal that are not cooperative. At any rate, this possibility suggests that many seemingly goal-directed acts and acts that can be done by individuals for their own sakes will be best understood as irreducibly cooperative, and thus that Bratman will have difficulty analyzing them in a non-circular way as well.

Though his remarks about it are brief, Bratman does attempt to sketch a strategy for answering this objection. He begins by drawing an analogy with a problem from individual action theory. To intentionally raise one’s hand, Bratman says, one’s hand must rise and the movement must be explained by certain attitudes that the agent has (desires, beliefs, intentions, and so on). But there might be cases in which the agent intends not just to raise her hand, but to do so intentionally. This raises a worry about circularity, since the object of the intention is *intentional* action. Bratman’s suggestion about how to avoid such circularity is very general. He says that we should try to understand cases of intending to perform intentional actions by looking first to the more “basic case” in which the agent simply intends to raise her hand (rather than to raise it intentionally). Then he suggests we might somehow extend our analysis of the basic case to the more complicated case in which the agent intends to raise her hand intentionally. So here is where Bratman draws his analogy with the problematic group action cases. He says we should examine, “basic cases in which each participant intends a joint activity understood in a way that is neutral with respect to shared intentionality.” We can then, “try to understand those cases in which the agent’s conception of the shared activity explicitly supposes it is a shared intentional activity, in a way that builds on our understanding of these basic

cases.”<sup>21</sup> Bratman then drops the issue, while admitting, “There will be hard issues about how exactly to go from basic cases to [cases like *dancing the tango*].”<sup>22</sup>

Obviously Bratman’s remarks are only meant to gesture at an approach to this problem.<sup>23</sup> There are nevertheless two reasons to worry about what he says. First, pertaining to his analogy with individual action, it isn’t obvious that we can simply analyze an instance of intentional arm-raising by characterizing the person’s arm’s rising in non-intentional terms and then identifying some attitudes that cause the movement. Anscombe, for instance, denies that we can do this. I will discuss this point in greater detail in the next section, but it is worth flagging here that Bratman’s analogy rests on a controversial assumption about intentional action. Second, even if one accepts that we can somehow abstract away the intentionality from an intentional arm-raising, it is not at all obvious why this should give one confidence that we can abstract away the cooperativeness or the “groupishness” from *dancing the tango*. It is at least somewhat plausible to suggest that we have a bodily movement remaining when we subtract one’s intention from an intentional arm-raising. It is unclear to me what we could plausibly claim remains when we subtract cooperation from *dancing the tango*.

To summarize, we saw that Bratman’s account is circular when used to analyze certain kinds of group actions—at least two of his necessary conditions for group action fall prey to this

---

<sup>21</sup> Bratman, “I Intend that We *J*,” in *Faces of Intention*, pp. 142-161 (quotation from pp. 147-148).

<sup>22</sup> *Ibid.*, p. 148.

<sup>23</sup> Building on Bratman’s suggestion, Christopher Kutz proposes a more worked-out solution. (See his, “Acting Together,” *Philosophy and Phenomenological Research* 61 (2000), pp. 1-31, (esp. pp. 14-15)). The proposal appeals to a participatory intention, whose content includes the agent’s playing her role in *dancing the tango*. “Dancing the tango” is, here, understood to be irreducibly cooperative. Kutz believes that he solves the problem by giving a genealogical account of how these intentions with irreducibly cooperative content bootstrap their way into existence, using only individual non-participatory intentions. However, regardless of how the agent comes to have intentions with irreducibly cooperative acts as part of their content, any analysis that makes use of such intentions will still appeal to the concept “group action.” So it isn’t clear how Kutz escapes the circularity charge. Petersson makes a similar criticism of Kutz’s argument. He suggests that what Kutz manages to do, through his bootstrapping account, is avoid a certain kind of *explanatory* circularity. That is, he explains how an agent can come to have participatory intentions without appealing to the notion of a cooperative act. But he and Bratman need to avoid circularity in their *analysis* of the concept, and Kutz’s argument fails to do this. (See Petersson, “Collectivity and Circularity,” pp. 142-143.)

circularity objection. Next, we saw that some of the cases that Bratman seemed to be in the best position to handle—cases like *painting the house*—might also resist a non-circular analysis, depending upon how we understand them. Then we saw that Bratman gives only a very general suggestion about how he might avoid circularity, and that there is reason to be skeptical about that suggestion. Finally, we saw that even Bratman alludes to the fact that we will encounter significant difficulties in trying to employ his proposed strategy. What lesson should we draw from all this? While no argument could prove this decisively, the foregoing certainly lends credence to the hypothesis that giving a non-circular analysis of the concept is impossible. Bratman’s attempts to do so seem to appeal quietly, but directly, to the concept “group action” itself. This is exactly what we would expect to happen were we trying to reduce a concept by appealing to features that are essentially parts of the whole that they constitute. This is exactly what we would expect to happen were we treating a non-accidental species as if it were an accidental one. (It is, for instance, what Ford thinks happens if one tries to define a particular species of animal by appealing to its organs and appendages).<sup>24</sup>

#### 4. Group Intentions

Earlier I noted that group action theorists often divide the field into two topics: “group action” on the one hand, and “group intention” on the other. One might wonder whether we could give an analysis of group action by appealing to group intention by treating group action as an accidental species of the genus “action,” where the relationship between the genus and the species is mediated by a group intention, understood as an accident that is responsible for bringing about the behavioral episode or action. Returning to Searle’s example, we might distinguish the case in which the people each individually try to escape the rain from the case in which the people perform a group action by appealing to the fact that the movements in the

---

<sup>24</sup> To give an example of similar kind of argument, in “The Representation of Life,” Michael Thompson considers various attempts to define “life-form.” All of them end up appealing to the concept “life-form” itself, which leads Thompson to conclude that it cannot be given a proper non-circular definition.

latter case stand in a particular kind of causal relationship to a group intention, whereas the movements in the former case do not.<sup>25</sup> In this section, I will argue against this strategy.

Before I begin, I want to mention a few examples of views about what group intentions are, in order to illustrate the range of views that my argument is intended to address.<sup>26</sup> (I will argue that none of them can be used to provide an analysis of the sort sketched above.) The accounts can be usefully divided into two groups: those on which a group intention is a state of mind of an individual, and those on which a group intention is some other state of affairs. Proponents of the former kind of view tend to talk of *individuals* having group intentions, whereas proponents of the latter kind of view tend to talk of *groups themselves* having intentions (or of a group of people sharing a single intention). Searle is someone who takes a group intention to be a state of mind of an individual—one that is irreducibly different from an individual intention.<sup>27</sup> Christopher Kutz similarly thinks that group intentions reside with the individual.<sup>28</sup> But unlike Searle, Kutz understands a group intention to be a mental state that is differentiated from an individual intention by its participatory *content*, and not by its form.<sup>29</sup> On the other side, Bratman does *not* understand a group intention to be a mental state of an individual.<sup>30</sup> He says that a group intention consists of, “a public, interlocking web of intentions

---

<sup>25</sup> Indeed, this is what Searle himself proposes.

<sup>26</sup> My use of the term “group intention” here is meant to include what has variously been called “we-intention,” “collective intention,” “shared intention,” “participatory intention,” and many other things as well.

<sup>27</sup> Searle’s account bears some similarity to Wilfrid Sellars’ view in his “Imperatives, Intentions, and the Logic of ‘Ought,’” in *Morality and the Language of Conduct*, eds. Hector-Neri Castaneda and George Nakhnikian (Detroit: Wayne State University Press, 1963), pp. 159–214.

<sup>28</sup> See his “Acting Together.”

<sup>29</sup> Raimo Tuomela and Kaarlo Miller argue for a view that is similar in some respects. (See their “We Intentions,” *Philosophical Studies* 53, (1988), pp. 367–389.)

<sup>30</sup> He does, however, allow that an agent can have an individual intention with participatory content. That is to say, he thinks that I can intend that “we” do something. But his view is that the individual can also share an intention with others, and thus that a group can have an intention.

of the individuals.”<sup>31</sup> The interlocking web then plays a role in bringing about activity on the part of the members.<sup>32</sup> David Velleman argues that a signed contract can itself constitute a group intention—one that exists outside of the minds of the parties to the agreement.<sup>33</sup> Finally, Margaret Gilbert believes that two people who jointly commit to acting together create a state of affairs in which they, as a collective, have an intention—one that persists even if each participant subsequently decides in her own mind that she will not actually carry it out.<sup>34</sup>

Obviously I am using “group intention” as a blanket term to cover a diverse set of concepts (or, perhaps, a diverse set of conceptions of what a group intention is). However, all of the aforementioned accounts have something in common. They all treat a group intention as a state of affairs that is in some way responsible for bringing about behavioral episodes or events that are themselves constitutive of group action. And so I think that *however* one conceives of a group intention, it will be tempting to analyze group action by appealing to the relationship that the behavioral episode bears to a group intention.

This strategy faces a challenge, however. The challenge is to specify *how*, exactly, the group intention must relate to the behavioral episode such that what occurs is a genuine group action. A similar challenge is familiar from individual action theory, and so I will quickly sketch the problem as it arises there. On one standard view, mental states such as beliefs, desires, and intentions are causally responsible for bringing about an individual’s intentional actions. But behavioral episodes are not intentional actions simply by virtue of their being causally related to such mental states. Imagine that I have an intention to confront my boss. This intention weighs heavily on me, causing me to ignore what I am doing. Absentmindedly, I knock over the cup of coffee sitting on my desk. My knocking over the cup does not become an intentional action

---

<sup>31</sup> “I intend that We *J*,” p. 143.

<sup>32</sup> *Ibid.*, p. 142.

<sup>33</sup> “How to Share an Intention,” *Philosophy and Phenomenological Research* 57 (1997), pp. 29-53.

<sup>34</sup> See her “Shared Intentions and Personal Intentions.”



simply by virtue of its being caused by my intention. Intentional actions, on this standard view, are behaviors caused by mental states that *rationalize* them. As one author puts it, intentional action is, “behavior [that] eventuates not only *from causes* but *for reasons*.”<sup>35</sup> So if, for instance, I desire a beer, and I believe that there is a beer in the kitchen, this belief-desire pair might very well cause me to go to the kitchen. Importantly, however, if I am asked to give my reason for going to the kitchen, I can cite that very same belief-desire pair. Nothing like this is true in the coffee cup case, because there is no rational connection between my intention to confront my boss and my knocking over the cup of coffee. Now for the difficulty. Donald Davidson imagines a mountain climber who desires to relieve herself of the burden of having a climbing partner, and who believes that releasing the rope will cause her partner to fall to her death. The presence of this belief-desire pair so disturbs the climber that her hands tremble, causing her to release the rope. As Davidson explains: “Beliefs and desires that would rationalize an action if they caused it in the *right* way—through a course of practical reasoning, as we might try saying—may cause it in other ways.”<sup>36</sup> In the mountain climber case, her releasing the rope *does not* appear to be an intentional action, because the belief and desire don’t appear to be causally related to the climber’s movements in the right way. This, despite the fact that the contents of the belief and the desire are of the right sort to rationalize releasing the rope. This is an example of what are called “deviant causal chains.”

The deviant causal chains problem arises for accounts of group action in an exactly analogous form. Imagine that the mountain climber is part of a foursome, three of whom conspire to kill the fourth. The climber has one of Searle’s group intentions (a mental state), which, if all goes well, will cause her to release the rope, sending the fourth climber to her death (while, say, the others stand guard to ensure there are no witnesses). But the mere presence of

---

<sup>35</sup> David Velleman, *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000), p. 5.

<sup>36</sup> “Freedom to Act,” in *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), pp. 63-82 (quotation from p. 79).

the group intention makes the climber nervous, causing her to tremble and to release the rope. In such a case, the group intention does *not* seem to be causally related to the rope-release in a way that allows us to say that the group intentionally acts, or that the member did the deed intentionally. The same problem will arise, I believe, no matter how we conceive of a group intention. Imagine that three climbers share an intention to kill the fourth—an intention that consists of one of Bratman’s interlocking webs of individual attitudes. The web consists, among other things, of one’s intending that the group kill the fourth, one’s knowing that the others also intend that the group kill the fourth, and one’s intention to “mesh” various aspects of one’s plan for doing the deed with the others’ plans. On Bratman’s view, all of this normally helps bring about the action by facilitating coordination. But imagine that on this occasion the whole web simply makes one of the members nervous, causing her to tremble and to release the rope. Once again, it does not seem like the intention relates to the deed in the right way.

There is a related problem, unique to group action, that I also wish to point out. Imagine that I have one of Searle’s group intentions, directing me to help kill a member of the climbing foursome. I have been conditioned by my parents (or my therapist) always to immediately and mindlessly form an individual intention to act in ways identical to those required by any group intention that I have. So, my group intention to kill the climber causes me to form an individual intention—a completely different kind of intention—to release the rope. Then my individual intention causes me, in the normal way, to release the rope. There is no temptation, here, to deny that I act intentionally. But I think that there is still a question about whether to understand what happens to be an intentional *group* action. I think we should question whether *just any* way that a group intention manages to cause an intentional action suffices for group action.<sup>37</sup> To see why, consider that any number of things might be causally responsible for the

---

<sup>37</sup> One might insist that in order for a group action to occur, the group intention must be the *proximate* cause of the bodily movement. This strategy has been proposed as a solution to the deviant causal chains problem in individual action theory. (See Alfred Mele, *Springs of Action: Understanding Intentional Behavior* (Oxford: Oxford University Press, 1992).) This strategy faces a notorious objection: when action

formation of an individual intention, but that we generally do not, on the basis of the existence of such causes, posit an entirely new form of action. For instance, my looking at my reflection in the mirror might always cause me to form an individual intention to diet. But the fact that seeing my reflection causes me to form an individual intention and then to act on that intention does not tempt us to think that when the action is carried out, it is anything more or anything less than a garden-variety case of individual action. If this is right, then why should the fact that a group intention is causally responsible for an intentional action in the way mentioned above prompt us to think that the action is anything more or anything less than a garden-variety case of individual action? I think it should not, even if the content of the group intention is otherwise of the right sort to rationalize what is ultimately done. The group intention has to cause the episode *in the right way*. And this, of course, raises the question: which way is the right way? I believe that however one conceives of a group intention—as something that resides inside the mind or outside of it, as something distinguished from an individual intention by its content or by its form—one must eventually face this question if one hopes to appeal to group intentions to give a reductive account of group action.

What I have argued, then, is that an account that analyzes group action in terms of behavioral episodes (or actions) and a relation to group intentions will face two pressing questions. One question is about the right way for an intention to relate to a behavior if the behavior is to be deemed an *intentional* action at all. The other question is about the right way for an intention to relate to the action if it is to be deemed an intentional *group* action. I think there is reason to be skeptical whether these questions can be given satisfactory answers. The

---

occurs, various physiological states must intervene between the intention and the behavior, calling into doubt whether an intention can actually be a proximate cause of behavior. Mele has a response to this objection (see esp. p. 202), but it remains controversial. (For a criticism from an opponent of the causal theory of action, see Scott Sehon, “Deviant Causal Chains and the Irreducibility of Teleological Explanation,” *Pacific Philosophical Quarterly* 78 (1997), pp. 195-213. For a criticism from a friend of the causal theory of action, directed at an argument similar to Mele’s, see John Bishop, *Natural Agency: An Essay on the Causal Theory of Action* (Cambridge, UK: Cambridge University Press, 1990), pp. 138-140.) I cannot resolve the debate here, but will just say that the reasons for doubting that appealing to the notion of a proximate cause will work for individual action theory also make me doubt that we can appeal to it in the case of group action.

deviant causal chains problem has persistently plagued causal theories of individual action. A prodigious literature has sprung up, attempting to specify what exactly “causing behavior in the right way” consists in. I cannot, here, review all the various proposed solutions.<sup>38</sup> But no solution has earned widespread acceptance, and there seems to be little agreement about which approaches to the problem are most promising. And if I am right, the situation for group action is even worse. Group action theorists face the exact same problem that individual action theorists do, as well as an additional problem related to deviant causal chains of a different sort. Thus, the prospects for a successful analysis of group action that appeals to group intentions causing or bringing about bodily movements appear grim.

It is thus reasonable to ask whether we have been led into this predicament by taking a fundamentally mistaken approach to the topic. And here we find a motivation for Anscombe’s suggestion that we not treat an intention as an extra ingredient of an event that the event itself can either have or lack. Returning for a moment to the individual case, the nervous rope-releaser undergoes movements that appear identical to those that would occur were she to release the rope intentionally. The apparently identical movements would constitute different events, however—one would constitute an intentional action, and the other would constitute a nervous response. It is an open question whether the differences between the two events can be characterized without appealing to the notion of “intentional action” itself. The appeal to causally efficacious mental states is, of course, an attempt to do so. But this is what invites the problematic question about how the mental states must cause the event. The deviant causal

---

<sup>38</sup> George Wilson gives the most famous sustained treatment of the problem in his *The Intentionality of Human Action* (Stanford: Stanford University Press, 1989). He concludes that the problem is devastating to the entire causal theorist’s project. But there have since been additional attempts to save the causal theory. Two of the best include John Bishop, *Natural Agency*, and Alfred Mele, *Springs of Action*. The former author suggests that an intentional action occurs when one’s behavior exhibits a certain sort of sensitivity to the content of one’s intentions, and the latter author suggests (as I mentioned in an earlier footnote) that intentional action occurs only if the intention is the *proximate* cause of the behavior. Still, these solutions remain controversial, and many have argued that the deviant causal chain problem can be generated for them as well.

chain problem arose because we treated the action's intentionality as something bestowed upon it through a causal connection to a mental state.<sup>39</sup>

Turning back to group action, I want similarly to question whether the deviant causal chains problems arise because the attempt to analyze group action by appealing to group intention constitutes a fundamentally mistaken approach to the topic. In the case of the group intention that causes the climber to tremble, we can say the same thing that we did in the individual case. The difference between a nervous response and an intentional action cannot be cashed out in terms of some extra ingredient that causes one but not the other. But now let us consider the case in which the group intention causes the climber to form an individual intention. I propose that the difference between an individual intentional action and a group action cannot be cashed out in terms of some extra ingredient (or ingredients) that brings about one but not the other. In the case of the climber who has been conditioned by her parents to form an individual intention when she has a group intention, she moves in ways that appear identical to the ways that she would were she to engage in group action. However, as I argued earlier, the apparently identical movements constitute different events—one constitutes an individual action, and the other constitutes a group action. Just as before, it is an open question whether the differences between the two events can be characterized without appealing to the notion of group action itself. The appeal to group intentions is an attempt to do so. But it is this strategy that invites the question about how the group intention must relate to the event in the first place. Here the deviant causal chain problem arises because we treat the intentional action's "groupishness" as something that is bestowed upon it through some kind of relation to a state of affairs that we call a group intention.

---

<sup>39</sup> Versions of this argument are advanced by Anscombe in "Practical Inference," in *Virtues and Reasons*, eds. R. Hursthouse, G. Lawrence, and W. Quinn (Oxford: Clarendon Press, 1998), pp. 1-34, and Candace Vogler, "Anscombe on Practical Inference," in *Varieties of Practical Reasoning*, ed. Elijah Millgram (Cambridge, Mass.: MIT Press, 2001), pp. 437-464 (esp. pp. 442-443).

These points can perhaps be clarified a bit by using the determinable-determinate terminology. The suggestion is that “event” is determinable. An intentional action and a nervous response are both determinates. It is a mistake to ask what additional accidental feature (like a causal relation to a mental state) differentiates the two. I am suggesting that the genus “action” is determinable as well. Individual action and group action are both determinates. It is a mistake to ask what additional accidental feature (like a causal relation to a mental state or to a web of individual intentions) differentiates the two, just as it is a mistake to ask what additional accidental feature differentiates an intentional action from a nervous response. This suggestion is offered as a diagnosis of why certain ways of attempting to reduce group action run into the deviant causal chains problems. The diagnosis also happens to support my hypothesis that group action cannot be given an analysis in which a group intention is treated as an accident.

### 5. Toward an Account of Group Action

If one accepts that group action is a non-accidental species of the genus “action,” one might then proceed in a way similar to Bratman, only with the assumption that some of the features that differentiate “group action” from others in the genus can only be defined or specified by appealing to group action itself. This is basically what Margaret Gilbert does.<sup>40</sup> Gilbert has tended to characterize her view as non-reductionistic, and so is not troubled if some of the features that she says are definitive of group action are, in the end, defined in terms of group action. However, what I am going to do here instead is sketch a positive account of what group action is by trying to explicate the sense in which it is a non-accidental species of the genus “action.”

Anscombe, as I noted earlier, takes it to be the mark of an intentional action that a particular kind of “Why?” question has application. As I understand her, the “Why?” question is a request to rationalize what is being done—it is a request to explain what is happening by

---

<sup>40</sup> Gilbert, *On Social Facts*.

appealing to reasons. According to Anscombe, when one responds to the “Why?” question in the spirit in which it is offered, one expresses an intention. So we can understand what she means by “intention” functionally. Whatever it is, an intention will be something suitable to be offered up when giving a rationalizing explanation. Anscombe’s purpose in *Intention* is to inquire into what sort of thing could serve this function.<sup>41</sup>

Richard Moran and Martin Stone provide a helpful illustration of what Anscombe has in mind, writing that she conceives of an intention as a, “location on a spectrum of unfolding action that fits an event to be described using concepts of intention in the first place.”<sup>42</sup> Here is how they describe that spectrum:

Picturing a line, on the far left will be found pure intentions, defined as cases in which the agent intends to do something but hasn't yet done anything else in order to do that. Moving rightward, the agent has more worldly deeds to show for his intentions: if he is described as  $\phi$ -ing or as intending to  $\phi$ , then, at this point, it will be correct to say that he is also doing various things in order to  $\phi$ , or because he intends (wants, aims) to  $\phi$ . At the far right, his performance is fully unfolded and finds description in the past tense: ‘He  $\phi$ -ed’ (or ‘He has  $\phi$ -ed’) will now be true, and not merely—what holds anywhere between these end points—‘He was  $\phi$ -ing.’<sup>43</sup>

So, if I am stepping on an airplane, and you ask why I am doing that, I might reply that I am traveling to Italy. (To make it clearer that this is a location on the spectrum, I might reply that I

---

<sup>41</sup> So understood, once one gives an adequate account of what an intention is, there will be no worry about deviant causal chains. The deviant causal chains problem arose for accounts that proceeded by giving an account of various mental states on the one hand (desires, beliefs, intentions, and so on), an account of behavior on the other, and then asking how the former could rationalize the latter. For Anscombe, an inquiry into what intentions are *simply* is an inquiry into what sort of item could rationalize action.

<sup>42</sup> Richard Moran and Martin Stone, “Anscombe on Expression of Intention: An Exegesis,” in *Essays on Anscombe’s Intention*, pp. 33-75 (quotation from p. 49).

<sup>43</sup> *Ibid.*, pp. 47-48.

am “stepping on the airplane in order to travel to Italy.”) That is, I express the intention that I have in getting on the plane, and in so doing locate myself somewhere in the middle of a spectrum representing the unfolding intentional action. Or, if you ask me why I am stuffing cash under my mattress, I might reply that I am travelling to Italy next summer. That is, I express an intention to do something in the future, and in so doing locate myself on the far left-hand pole of a spectrum representing the unfolding intentional action. Of course, these seem to be expressions of two different things. In the former case, I express the intention with which I do something (what philosophers sometimes call an “intention in action”), whereas in the latter case I express an intention that seems to be, in Anscombe’s words, “a purely interior thing” (or what other philosophers sometimes call a “future-directed intention”).<sup>44</sup> However, Moran and Stone’s purpose in calling our attention to the spectrum is to reveal the unity of two seemingly distinct concepts. Both are locations on the spectrum—a spectrum representing an ongoing action as it unfolds.

In both instances, the expression of intention explains what is happening. But it does not do so merely by showing that what is happening eventuates *from causes* of some kind or another. It does not, for instance, report on neural firings or brain activity. By locating what is happening on the spectrum, the expression explains what is happening by appealing to *reasons*. The idea here is that ongoing actions that take up time are event-processes that exhibit a rational structure. That is to say, the action that the spectrum represents is capable of rationalizing its sub-parts. This is why locating oneself at a particular place on the spectrum both signals that one is acting for reasons and reveals what those reasons are.<sup>45</sup> For instance, the single ongoing action *my travelling to Italy* rationalizes *my driving to the airport, my printing*

---

<sup>44</sup> Anscombe, *Intention*, p. 9.

<sup>45</sup> The thesis that actions rationalize their sub-parts is defended at length by Michael Thompson in “Naïve Action Theory,” in *Life and Action*, pp. 85-146.



*a boarding pass, and my stepping on an airplane.*<sup>46</sup> To say “I’m travelling to Italy” is to give my reason. Once we endorse this view, however, we see that an intention cannot be conceived as an extra ingredient added to a behavior or bodily movement, rendering it intentional. This is because an intention is defined as a location on a spectrum that represents an unfolding *intentional action*. The concept “intentional action” is thus prior to the concept “intention.”

If we endorse such an account of individual intention, what, then, can we say about group intention? As a kind of intention, I think we should also understand it functionally—its function being to enter into rationalizing explanations. A group intention, I propose, also accomplishes this, in part, by locating what an individual is doing along a spectrum that represents an unfolding action. Imagine a cymbal player who, upon entering Carnegie Hall, is confronted with Anscombe’s “Why?” question. She might answer: “We are playing Tchaikovsky’s *1812 Overture* tonight.” The answer is an expression of an intention, which locates her present doings at the left-hand pole of the spectrum—the orchestra has yet to begin its performance. Or, imagine the moment when the cymbal player crashes her cymbals. Confronted with Anscombe’s “Why?” question, she might similarly answer: “I am crashing my cymbals now because we are playing Tchaikovsky’s *1812 Overture*.” This expression of an intention locates her in the middle of the spectrum, at the place within the symphony where it is appropriate for her to play her instrument. In both cases, the intention plays a role in the rationalizing explanation of what is happening.

Thus far, my conception of a group intention parallels the Anscombian conception of an individual intention almost perfectly. Now I want to add a bit of complexity, which will differentiate group action from individual action. A group action unfolds not only over time, but synchronically across the group’s members. Here is what I mean. When a football team executes

---

<sup>46</sup> Michael Thompson begins “Naïve Action Theory” by speculating that, “the explanation of action as it appears most frequently in human thought and speech is *the explanation of one action in terms of another*.” Thompson’s examples include pulling a cord in order to start an engine, cutting wires in order to repair a short-circuit, and breaking eggs in order to make an omelet. (See p. 85.)

a pass play, it usually takes ten seconds or so. If I, as a spectator, arrive from the concession stand five seconds in, I might see just part of the play—perhaps the football is already in the air, and the receiver is already streaking downfield to catch it. I might see, that is, just a portion of the action’s *temporal unfolding*. As a spectator, I might also discover that there is too much action on the field to take in at any given moment. So I might home in on a particular player—say, the quarterback or the receiver—and ignore the rest. As was true after my tardy arrival from the concession stand, here I observe only a part of the pass play that the group performs. In this case, though, what I witness is not a portion of the action’s temporal unfolding, but instead a portion of the action’s unfolding as it manifests itself in the quarterback or in the receiver—a portion that unfolds at the same time as other portions, that are performed by other participants. If I want to see the whole action unfold along both dimensions, I must be present from start to finish, and I must take care to survey the entire field throughout.

Shifting perspectives now, if I am one of the players on the field, it will be insufficient to answer Anscombe’s “Why?” question by locating myself along the spectrum representing the temporal unfolding of the play. I need to locate myself within the action’s *synchronic unfolding* as well. That is to say, I need to locate myself within the set of *roles* assigned to players on the field at any given moment. As I understand them, “roles” are defined in terms of the group action. To be a quarterback, a receiver, a running back, or a tight end is to be responsible for carrying out certain portions of a particular kind of group action. That a group action apportions responsibility through the assignment of distinct roles is crucial to its ability to rationalize its sub-parts. Just as my traveling to Italy can rationalize my printing a boarding pass, my going through security, or my stepping onto the plane, depending upon where I am in the temporal progression of the action, so too can our executing a pass play rationalize my blocking, your getting out of the way, or Bob’s catching the ball, depending upon who occupies which role. So a group intention is a location along a spectrum representing an unfolding intentional action *and* a member of a set of inter-related roles necessary to carry out the group action (or, rather, it is

an ordered pair containing these two items). If this is right, a group intention is defined in terms of an intentional group action.

Through our discussion of group intention, a picture of what group action is has come into view. Like an individual action, a group action is a rationally structured process-event. It is constituted by temporal phases—sub-actions—that are rationalized by the whole.<sup>47</sup> But unlike an individual action, a group action is constituted by roles, and the sub-action characteristic of each role is also rationalized by the group action. *Our running a pass play* rationalizes the quarterback's throwing the ball, and the receiver's catching it. *Our conducting an interrogation* rationalizes the good cop's acting reassuringly, and the bad cop's acting tyrannically. Group actions are event-processes that rationalize their sub-parts in just this way. *This is what it is for something to be a group action.*

I will conclude with some remarks about how to differentiate group action from individual action. Imagine two individuals who undergo identical movements: they each drop back and throw a football. One does so by himself, as part of an individual action—he is, say, just dropping back and throwing for its own sake. The other does so on a field with others, as part of a group action. The two have something in common. Both ball-throwings are process-events that rationalize their sub-parts. That is to say, they are members of the same genus—they are both *actions*. But what the agent does as he releases the ball by himself is rationalized differently than is what the quarterback does as he releases the ball in the midst of a pass play. The intention with which the quarterback carrying out the play acts is not *merely* a location on our spectrum, but a location on our spectrum *and* a member within a set of roles defined by the action itself. This marks the difference between the otherwise seemingly identical actions.

So the difference between group action and individual action has to do with how each rationalizes its sub-parts. But if an action is understood to be a rationally structured process-

---

<sup>47</sup> Ben Laurence has also suggested that individual doings can be rationalized by a group action. (See his “An Anscombian Approach to Collective Action,” in *Essays on Anscombe's Intention*, pp. 270-294 (esp. p. p. 281).)

event that rationalizes its sub-parts, then this is just to say that group action and individual action differ in respect of *action* (just as green and red differ in respect of *color*). Group action and individual action are two distinct determinates. Each is a non-accidental species of the genus “action.”

While my examples in this chapter have tended to be of a smaller scale, I hope it is clear how this account could be applied to society-wide action. The people of the United States might decide to occupy Iraq. Doing so would be an action that would rationalize its subparts, which are carried out by individual citizens each with her own role. Some citizens might actually travel to the Middle East to be part of the military operation, if that is their role. There, if the soldiers are asked why they are travelling around the dessert, or are manning a security checkpoint, they can answer, “We are occupying Iraq.” At the same time, the citizens back at home might also do things that are rationalized by this activity. In the past, citizens might have rationed their gasoline or paid a war tax. That was their role in the effort, and if asked why they were doing these things, their answer might have been, “We are fighting a war.” Nowadays, a citizen is more likely to be told to “support the troops” or to “support our president” because “We are at war.” A citizen is supposed to see these sub-actions—these acts of support—as her contribution to the war effort. This is her role. Despite the scale being much larger, the basic story is the same as the one I have told over the course of this chapter. We now have an account of what it means for an entire society to act as a group, and we can now turn to my argument that societies are group agents.

## Chapter 2: Society as a Natural Agent

### 1. Introduction

Having explained what it means for an entire society to act together as a group, I will now argue that human societies are naturally rational agents, and that their members are naturally constituents. According to this view, no one needs to decide to transform a multitude living in a geographical region into an agent, and no one needs to consent to being part of the agent (as is the case when we speak of *artificial* group agents).<sup>1</sup> Whether it be a small, isolated city with mostly informal customs and practices, or a large nation with complex formal institutions, a human society *just is* an agent, and its members *just are* constituents.

As rational agents, I argue, human societies are under normative pressure to engage in action that is, quite generally, good or worthwhile. In this, they differ from many other types of groups. For instance, the fact that putting on a ballet, providing assistance to the poor, and cleaning up a national park are all things that are worth doing arguably puts no normative pressure on a corporation to do these things. Businesses, most would agree, exist to pursue profits, not to pursue worthwhile endeavors generally. So the mere fact that some project is worthwhile arguably does not give the corporation an even *prima facie* reason for pursuing it. My view is that human societies differ from businesses in this regard. They are not, as some believe, merely under normative pressure to pursue some particular end—such as realizing a state of affairs in which everyone’s property rights are protected—or to promote some particular value—such as freedom or physical security—irrespective of what it would be good or best to do

---

<sup>1</sup> Hobbes, for instance, talks of individuals creating an “artificial man,” capable of acting. This is accomplished through an agreement or decision on the part of the members. (See his *Leviathan*, ed. Richard Tuck (Cambridge, UK: Cambridge University Press, 1991), p. 9.) When I claim that society is a natural agent, I mean that no such agreement or decision is required.

on the whole. A society is under (defeasible) normative pressure to engage in worthwhile endeavors generally.<sup>2</sup>

The sorts of collective acts that societies are under normative pressure to perform require the use of state institutions. It is by using such institutions that a people is able to make decisions about what to do, and organize itself so as to act effectively. As such, I maintain that the state's proper function is not merely to create an environment within which individuals are free to carry on with their individual plans of life, engage in their own personal projects, and pursue their own conceptions of the good. The state's proper function includes helping to organize and carry out collective plans and projects.<sup>3</sup> This is an expansive view of the state's function, and my goal in this chapter is to defend it.

## 2. Natural Agency

Hegel is the philosopher most closely associated with the view that society is naturally an agent. But he attracts considerable skepticism by appealing to the notion of a collective spirit or consciousness, possessed by an entire nation.<sup>4</sup> While Hegel's precise meaning is a matter of controversy, political and legal thinkers writing in his wake routinely posited the existence of supernatural life-forces that supposedly animated collections of people.<sup>5</sup> The metaphysical extravagance of such views renders them dubious, and like most contemporary philosophers, I wish to distance myself from them. So my goal here is to argue that society is naturally an agent,

---

<sup>2</sup> I am not, here, claiming that the normative pressure to act collectively in ways that are good or worthwhile cannot be trumped by considerations of right. What I say in this chapter is consistent with the claim that right ultimately prohibits us from acting upon this normative pressure. My narrow aim, for the moment, is just to establish that there is defeasible or overridable normative pressure, impelling societies to, quite generally, do worthwhile things.

<sup>3</sup> In this sense, the view I argue for here is opposed to a political liberal view, like Rawls's. In this chapter, I will not argue against political liberalism directly (though I will offer some arguments against it in Chapter 6). Here I am simply giving a positive argument for a political view that I take to be a rival to political liberalism.

<sup>4</sup> G. W. F. Hegel, *Elements of the Philosophy of Right*, ed. A. Wood, trans. H. B. Nisbet (Cambridge, UK: Cambridge University Press, 1991), p. 312.

<sup>5</sup> For a nice, brief overview of such views, see Christian List and Philip Pettit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011), pp. 8-10.

while maintaining a commitment to naturalism.<sup>6</sup> I will begin in this section by offering a naturalistic Aristotelian account of agency, and then argue that low-level, non-rational agency may be properly ascribed to groups of non-human animals. In subsequent sections, I will build on this account, and argue that we can plausibly ascribe *rational* agency to human societies, and that society's members, as constituents of the group agent, are under normative pressure to participate in some of the society's collective activity.

When a particular object moves or changes, we sometimes attribute the movement or change to the object itself, rather than to external forces acting upon it. The precise nature of the attributions will vary, depending upon the type of object under consideration—our account of how the movements of a sponge are attributable to it will differ from our account of how the movements of a rational human are attributable to her. But at the most general level, we make such attributions in instances where there occurs *self*-movement or *self*-change, as opposed to movement or change initiated from without. To say that a thing moves itself, and thus that its movements are attributable to it, is to ascribe agency (in a broad sense) to it.<sup>7</sup> This Aristotelian distinction between self-movement and movement externally imposed is most apparent when considering animal movement. If a particular flying squirrel sees an acorn, scurries to the edge of a branch, and then glides to the adjacent tree to collect it, we say that she was moving under her own power. We say that she *moved herself*. If, on the other hand, our squirrel is lifted up and is thrown about by hurricane winds, the movement is explained by scientific laws governing the effects of the wind on physical objects. Just as a similarly sized rock or a branch would be, the squirrel is acted upon from without. Her journey "happens to her"; it is something she "undergoes."<sup>8</sup> In addition to animals, Aristotle tells us that plants, "possess in themselves an

---

<sup>6</sup> In saying I am committed to naturalism, I mean merely that I do not want to posit the existence of supernatural entities or anything that conflicts with modern science.

<sup>7</sup> The broad sense of "agency" I have in mind extends to non-rational life-forms.

<sup>8</sup> Aristotle himself uses the example of a human being carried by the wind. (See Aristotle, *Nicomachean Ethics*, 1110<sup>a</sup> 1-5.)

*originative power* through which they increase or decrease in all spatial directions."<sup>9</sup> When a particular plant grows a leaf or a flower, we judge the plant itself to be the source of the change. On the other hand, if a plant changes because it is eaten or because it is burned by a forest fire, we judge that it has simply been acted upon by external forces.

This distinction between what an entity *does* on the one hand, and what *happens to it* on the other, is rather fundamental to our understanding of the world. Without it, not only would we have difficulty understanding what it is for a plant or an animal to carry out life-processes, but we would have difficulty understanding what it is for a human to do something for which she is responsible. The Aristotelian thesis that I advocate is that attributions of self-movement and self-change to an object presuppose that the object in question bears a specific kind of form. Movement or change is attributable to a particular object when either is explained by appealing to the fact that the object is functioning in a way characteristic of bearers of its form.<sup>10</sup> So when one asks why a particular flying squirrel is gliding toward an acorn, or away from a predator, some appropriate replies might include, "That is how flying squirrels nourish themselves," or, "That is how flying squirrels protect themselves," or even, "That is just what flying squirrels do."<sup>11</sup> These replies appeal to the form that the creature bears—the form of a squirrel—in order to explain its particular movements. The Aristotelian attributability thesis is that by virtue of their emanating from the creature's squirrel-nature, the movements are attributable to it. That is to say, the squirrel is exercising agency.

The presupposition that there is something that it is for a squirrel to function properly will likely meet with resistance, as the notion of living beings having a "function" can seem more

---

<sup>9</sup> Aristotle, *On the Soul*, trans. J. A. Smith, in *The Complete Works of Aristotle, Volume One*, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984), p. 658 (413<sup>a</sup> 25-30), my italics.

<sup>10</sup> See Aristotle, *Physics*, (192<sup>b</sup> 10- 194<sup>b</sup> 15).

<sup>11</sup> Of course, one might also answer by mentioning that certain neurons are firing in the squirrel's brain, or that particular molecules are interacting in a way such as to produce this larger movement. But at least for some versions of the "why?" question, such a response would miss the mark. For instance, when a small child asks why the squirrel glides from branch to branch, it would seem non-responsive to say something about the laws of physics, rather than something about how squirrels go about living their lives.



at home with pre-Darwinian conceptions of the universe than with our own. We generally think of artifacts designed by intelligent beings for specific purposes—knives, clocks, cars, and so on—to be paradigmatic examples of objects that can be properly evaluated in terms of whether they are functioning well or poorly. Animals and humans are not designed for purposes in the way that clocks or cars are, and for this reason we might think it is misguided to talk about their "function." If my appeal, here, to "function" is indeed problematic, then the Aristotelian attributability thesis rests upon an objectionable metaphysics. But I deny that this appeal to "function" is objectionable, at least once it is properly understood. To say that S is a bearer of a particular form, and that there is, as such, something that it is for it to function properly, needn't imply that S has a function in the sense of "being designed to serve some end that its creator had in mind." It might simply imply that S characteristically does certain things, and has a characteristic way of doing what it does. Indeed, some have read Aristotle in exactly this way. Here is Christine Korsgaard:

[A]ccording to Aristotle, a living thing does have a definite purpose, in the sense of a "what it does." That purpose is to keep its own form, its own manner of functioning, in existence. It does this in two ways: first, through the continuous self-rebuilding activities of nutrition, which maintain its form in a spatio-temporally continuous stream of matter, and, second, through reproduction, by which it imposes its form on individually distinct entities. This is not a controversial metaphysical thesis about what living things are for, but rather a definition of 'living.' So far as this goes, there is nothing objectionable about Aristotle's teleology.<sup>12</sup>

---

<sup>12</sup> "Aristotle's Function Argument," in her *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (Oxford: Oxford University Press, 2008), pp. 129-150 (quotation from p. 141).

Thus, we needn't treat with suspicion the suggestion that organisms can "function properly" or "malfunction." Indeed, Michael Thompson has argued at length that our familiar thoughts about living things generally, and about particular species specifically, are saturated with judgments about their proper constitution and about what it is for them to function properly—judgments that in no way presuppose that living things are designed with some further purpose in mind.<sup>13</sup> Such judgments, Thompson says, are familiar to anyone who has read wildlife field guides or who has watched nature programs on public television, both of which make statements like, "The Texas bluebonnet harbors nitrogen-fixing microbes in certain nodes on its roots," or, "The yellow finch breeds in the spring, attracting its mate with such and such song."<sup>14</sup> The general form of such judgments, Thompson says, is: "S's characteristically (or typically) are/have/do F," or, "This is (part of) how S's live: they are/have/do F."<sup>15</sup> Thompson calls these judgments "natural-historical judgments."

Natural-historical judgments constitute a unique "form of thought," and according to Thompson, they are not reducible to other kinds of judgments. For instance, they are not reducible to statistical judgments, as "S's characteristically have F" does not imply that all of them do. Cats characteristically have four legs, but most of us have encountered a cat with three. "S's characteristically do F" does not even imply that most do.<sup>16</sup> Our nature program might report that, "The mayfly characteristically mates in early spring," despite the fact that only a small portion of them succeed in doing so before dying off. Natural-historical judgments also

---

<sup>13</sup> Thompson also insists that *even if* God created an organism with some further purpose in mind, these kinds of judgments would be unaffected. He writes, "[E]ven if the Divine Mind *were* to bring a certain life-form into being 'with a view to' securing an abundance of pink fur along the shores of the Monongahela, this 'purpose' would have no effect on the inner natural teleological description of that form of life, for this is its inner causal structure, taken generally." (See his "The Representation of Life," in *Life and Action: Elementary Structures of Practice and Practical Thought* (Cambridge, Mass.: Harvard University Press, 2008), pp. 25-84 (quotation from p. 79).)

<sup>14</sup> Ibid., pp. 64-65.

<sup>15</sup> Ibid.

<sup>16</sup> Ibid., p. 68.

cannot be analyzed by appealing to what an S will do if it is not interfered with, or what it will do under "normal conditions." To say that "S's always do F so long as they are not interfered with" presupposes a conception of "interference" that can only be spelled out by appealing to the natural-historical judgment itself.<sup>17</sup> We only know, for instance, whether a particular mother's nursing her cub counts as interference if we know whether it is true that, "The baby S is nursed by her mother" (understood as a natural-historical judgment). Finally, Thompson also denies that natural-historical judgments can be reduced to some other species of normative judgment.<sup>18</sup> He admits that natural-historical judgments are normative (or, at least that they resemble normative judgments) in the sense that we may infer that, "This S is defective in that it is not F," from, "The S is F" (understood, again, as a natural-historical judgment), and, "This S is not F."<sup>19</sup> But the sense of "defect" in play here cannot be analyzed by appealing to some other normative standards that are themselves independent of the natural-historical judgments.

Let me now explicitly connect the thesis that movements or changes are attributable to particular entities with the observation that the entities, as bearers of a particular form, are proper objects about which to make natural-historical judgments. Natural-historical judgments appear to utilize something similar to the notion of attribution. "S's characteristically do F" attributes F to S in a sense similar to the one we are considering. Thompson does *not* attempt to further reduce these judgments by, for instance, defining independently "S" and "doing F" and then giving an account of what it would mean to attribute F to S. In fact, if I understand him correctly, he thinks it would be a mistake to attempt to do so, as he believes the notion of "doing F" will be conceptually dependent on "S." (He argues, for instance, that "eating" or "growing a leaf" cannot be defined without mentioning the kinds of organisms that do these things, as each is essentially part of the life cycle of the organism itself.) Whether or not Thompson is correct

---

<sup>17</sup> Ibid., p. 70.

<sup>18</sup> Ibid., pp. 73-76.

<sup>19</sup> Ibid., p. 80.

about this, if we think that we grasp what it is to make a natural-historical judgment—if we understand the form of judgment being made in the wildlife field guides and in the nature programs—we can then appeal to these judgments in order to explain *particular* movements of *particular* individuals. For instance, we judge that, "Plants engage in photosynthesis for the purposes of sustenance and growth." So long as we endorse the content of this natural-historical judgment, and grasp its (perhaps irreducible) form, we can judge that a particular movement of a particular plant is constitutive of a process that plants characteristically initiate or carry out. So we can judge of a particular plant that it is changing or moving *itself* when we observe it absorbing energy from the sun.<sup>20</sup>

My discussion, here, about attribution of movement or change to individual living beings has been offered in preparation for a discussion about the possibility of attributing movement or change to groups—the possibility, that is, of ascribing agency to them. So let us consider whether we can plausibly make natural-historical judgments about a group of individual organisms, such as an ant colony. To make such judgments, one must, of course, presuppose that the colony bears a particular form, with a characteristic way of functioning. But this is a perfectly natural way of understanding what an ant colony is—it is a self-maintaining entity, which has characteristic ways of carrying out the self-maintenance. At any rate, that is what wild life field guides and nature programs apparently assume when they say things like, "The ant colony characteristically maintains a system of tunnels underneath the ground," or, "This is how an ant colony feeds its members: it stores food throughout the mound of dirt, and then distributes it as needed." Such judgments bear all the marks of a natural-historical judgment. They are not statistical. Certainly to say that ant colonies characteristically maintain a system of tunnels

---

<sup>20</sup> Philippa Foot, expounding upon the relationship between these judgments and our understanding of what particular organisms are doing, writes, "without [referencing natural-historical judgments], life-activities such as eating or reproducing could not even be identified in an individual. Eating, for instance, is essentially, conceptually related to nourishment, and could not be conclusively identified by a story about the taking in, crushing, transforming, and spewing out of substances since, for all that, its purpose might not be the maintenance of tissue but, say, skunk-like defense." (See *Natural Goodness* (Oxford: Oxford University Press, 2001), p. 28.)

underneath the ground does not imply that *all* of them do. Any child who has tried to keep an ant farm knows that some colonies fail. Furthermore, there is no inconsistency in saying that, “This is how an ant colony feeds its members: it stores food throughout the mound of dirt, and then distributes it as needed,” but then admitting that most of them fail at this, and claiming that this is why so few colonies actually survive for very long. Finally, the judgments are also normative, in the same irreducible sense that our judgments about individual ants are. It isn’t as if the failure on the part of a colony to maintain tunnels is a defect merely according to a set of normative standards independent of these purported natural-historical judgments—for instance, it isn’t as if the ant colony is defective because its failure undermines the interests of the child who wants to observe ants burrowing tunnels in the ant farm that she purchased. It is normative just in the sense that we may infer, “This ant colony is defective in that it is not F,” from, “The ant colony is F” (understood as a natural-historical judgment), and, “This ant colony is not F.”

To drive home the point that we generally do consider an ant colony to be a bearer of a particular form, and to be capable of functioning well or poorly, let us contrast it with a group that bears no such form. Consider a menagerie composed of random animals—lions, lynxes, porcupines, and owls, let us say. While we could certainly produce statistical generalizations about what will go on when we place similarly constituted groups of animals together, we would not have any way of determining which behaviors constituted “functioning well” and which constituted “functioning poorly.” The problem here is not merely epistemic. The problem is that we don’t have a grasp on what it even means for a menagerie to function well. This is because, I submit, a menagerie is a “mere heap” of animals, and not a unity bearing a form.<sup>21</sup>

But, unlike with the menagerie, it *does* seem proper to make natural-historical judgments about some groups of organisms that we encounter in nature—colonies, packs, herds,

---

<sup>21</sup> “Mere heap” is Aristotle’s phrase, which he uses to draw a contrast with things that have unity. (See *Metaphysics*, 1045<sup>a</sup> 5-10.)

and so on. As such, we may attribute to the colony the maintenance of the tunnels and the distribution of the food—that is, we may ascribe agency to the entire group. The only worry one might have, I think, is that while making such judgments reveals no incoherence or conceptual mistake, we might be able to reduce them to judgments about individual members. So to say that, “The colony characteristically does F,” one might think, is just a way of saying something about what individual ants characteristically do. I am not going to argue against this possibility, however, because my ultimate purpose here is to apply this analysis to humans, and I have already argued (in Chapter 1) that group action on the part of humans cannot be reduced to individual action.

So we have, then, an account of what it would mean for a group to naturally constitute an agent. Crucially, this account is naturalistic. It requires no appeal to anything supernatural or metaphysically extravagant. There is no collective consciousness, or spirit of the colony. The view simply requires that we treat the ant colony as a bearer of a form with a characteristic way of functioning, just like an individual organism.

### 3. The Difference Rationality Makes

Rationality sets humans apart from plants and from other animals. A characteristic human life does not consist merely in self-maintenance through nutrition, nor does it consist merely in instinctual movement. It consists in exercising one's capacity for rational choice—it consists in acting on the basis of reasons.<sup>22</sup> So when an individual human moves, she often moves because she made a choice. In fact, she usually does not choose merely to engage in a movement, but to engage in a movement *in order to*  $\phi$ . For instance, an individual doesn't merely choose to move her arm, but moves it in order to get a glass of water, in order to pat her friend on the shoulder and give her reassurance, or in order to cast the deciding vote to send troops into combat. Of

---

<sup>22</sup> See Aristotle, *Nicomachean Ethics*, Book I and Book VI (esp. 1144<sup>a</sup> 5-10); Korsgaard, "Aristotle's Function Argument"; Foot, *Natural Goodness*, pp. 52-65.

course, humans engage in nutritive self-maintenance, as plants do, and they live lives involving sensation, as other animals do. So why single out rational activity, as Aristotle undoubtedly does, and identify the human function with the exercise of one's rational capacities? The answer, Korsgaard explains, has to do with how rationality alters significantly the *way* that a creature that possesses it lives its life:

The power of choice changes the way that we carry out the activities that we share with the other animals, such as housebuilding, childrearing, hunting or collecting food, playing, and sexual activity. Human beings approach these activities creatively and develop various ways of going about them among which we then choose. But we also do things that the other animals don't do at all, like tell jokes and paint pictures and engage in scientific research and philosophy. So rational choice introduces a whole new sense of *life*, a new sense in which a person can be said to "have a life."<sup>23</sup>

For this reason, we make a natural-historical judgment, of a sort: "The human being lives a life of rational activity."<sup>24</sup> That is, a being that functions as a human at all—a being that lives a *human* life—will move herself in ways that she judges herself to have reason to move. This is why when Aristotle speaks specifically of *human* self-movement or self-change, he is generally speaking of fully-fledged voluntary actions—like patting a friend on the shoulder to reassure her, or raising one's hand to send troops into combat.<sup>25</sup>

---

<sup>23</sup> Korsgaard, "Aristotle's Function Argument," p. 143.

<sup>24</sup> For an argument that we can sensibly apply to humans the patterns of natural normative judgment that Thompson discusses, see Philippa Foot, *Natural Goodness*, pp. 38-51. For related arguments from a Kantian perspective, see Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), esp. pp. 27-44.

<sup>25</sup> See esp. *Nicomachean Ethics*, Book III. At various places, Aristotle simply equates human life with action. See, for instance, *Politics* 1254<sup>a</sup> 5-10.

Aristotle thought, I believe, that being the sort of creature that functions by acting for reasons places the individual human under a requirement to aim to do what is *best*, or to aim at the highest good generally. But even if one does not agree with this, most can agree with the claim that to demonstrate that  $\phi$ -ing would be *good* puts some kind of pressure on a rational human being to  $\phi$ . That is, we think that by pointing out that  $\phi$ -ing is good or that it is superior to not  $\phi$ -ing, we are giving the rational agent a justificatory reason for  $\phi$ -ing.<sup>26</sup> Of course, this leaves unspecified what, exactly, an action's being good or worthwhile consists in. But it does suggest a general picture on which a human functions well when she identifies things that are, quite generally, worth doing with her life, and then does them. In fact, this is why when we speak of a "good human" or a "good human life," we generally speak of the quality of the individual's actions.<sup>27</sup> The good human is one who acts in ways that are good or worthwhile. This distinguishes a human from a plant or a non-rational animal. It also distinguishes her from artificial group agents that are created for a specific purpose—to return to the example from earlier, it distinguishes her from corporations whose overriding purpose is to maximize profits.

Groups of humans also possess this capacity for movement in accordance with reason. A group might deliberate and settle upon  $\phi$ -ing.  $\phi$ -ing might, in turn, rationalize certain sub-actions, carried out by the members. This is what I argued in Chapter 1. Given this, what should we say about human societies—by which I mean, roughly, groups of people that occupy a geographic region, who interact and who have contact with one another.<sup>28</sup> About all human societies, it ought to be acknowledged that they are entities with the capacity for rational

---

<sup>26</sup> For a discussion of Aristotle's views on this point, see Sarah Broadie, *Ethics With Aristotle* (Oxford: Oxford University Press, 1991), pp. 179-185.

<sup>27</sup> Foot makes this point, noting that it is peculiar that while a judgment about something's being a "good plant" is a judgment about the plant as a whole, when we make a judgment about something's being a "good person," we are concerned specifically with her rational will. (See her *Natural Goodness*, p. 66.)

<sup>28</sup> A "society" in this sense could be as small as a village with informal practices and institutions, or, in the modern world, as large the entire world population with its complex and formal practices and institutions. A "society" in this sense will also include conflicted or fractured societies in which there is little cooperation, despite cooperation being possible.



activity. A society might exercise this capacity by, for instance, deliberating about what to do, and settling upon "pursuing the common good," along with a way of going about doing this.<sup>29</sup> This group action will, in turn, rationalize its constitutive subparts—a citizen paying her taxes, a soldier defending the border, a public hospital employee distributing benefits, and so on—by assigning roles and apportioning responsibilities to members (in the way laid out in the previous chapter). Whether or not it actually performs this action, a human society possesses the capacity to do so (and to perform other actions).

The very same considerations that justify identifying the exercise of rational capacities with the human function also justify identifying the exercise of rational capacities with society's function. As it did with individual humans, the possession of this capacity impacts profoundly the way that human society goes about doing what it does. A human society, like the ant colony, engages in self-maintenance—a typical human society needs to maintain its infrastructure, institutions, and practices in order to persist and to continue to support human life and interaction. But it differs from the ant colony in that it can approach this endeavor creatively. The group can call into question, deliberate about, and alter the way that it performs this maintenance—for instance, by distributing roles and responsibilities so as to increase efficiency, or to meet standards of justice, or even to meet certain aesthetic standards.<sup>30</sup> Furthermore, a group's capacity to perform actions collectively introduces a new sense of what it means to "live a collective life," so to speak. For the ant colony, living a collective life means little more than engaging in self-maintenance, whereas for human societies, living a collective life often involves carrying out collective projects, thought to be worthwhile. Nations can and do commit themselves to space travel, to curing cancer, to artistic excellence, to academic achievement, to athletic accomplishment, to the realization of economic justice, to dominating a geographical

---

<sup>29</sup> I discuss what it means for the society to make a choice in Chapter 5.

<sup>30</sup> I mention aesthetics because societies often concern themselves with things like maintaining infrastructure that is not only useful, efficient, and economical, but also attractive or beautiful.

region, and to the achievement of the common good.<sup>31</sup> And so societies, like individual humans, function not by maintaining themselves instinctually, but both by choosing how they will maintain themselves, and by choosing additional activities that will constitute their collective life in the sense I outlined above. It seems natural, then, to identify society's function with the exercise of its rational capacities, just as we did with individual humans.

As we said earlier, natural rational agents with the capacity to engage in rational activity are under normative pressure to aim at doing things that are good or worthwhile. So, we should conclude that to demonstrate that  $\phi$ -ing would be good puts some kind of normative pressure on a society to  $\phi$ . Here too, I am leaving unspecified what "good action" or "worthwhile activity" would consist in. But the general picture is one on which a human society functions as it ought when it identifies things that are worth doing, and then does them. This gives us a standard for evaluating a human society, considered as such.

#### **4. The Relationship Between the Individual and the Group**

That a society can be a "good society" by acting well does not entail, in itself, that individuals ought to work to realize the good society. Reflective rational individuals, after all, can ask: "Why should I concern myself with whether my society functions as it ought?" Ultimately, I aim to demonstrate that the fact that a society ought to  $\phi$  places its constituents under normative pressure to cooperate in its efforts to  $\phi$ , or to do what they can to ensure that it  $\phi$ 's. But securing this conclusion requires investigating more closely the normative relationship between an individual agent and the group agent that she helps to constitute. In this section, I identify two ways that an individual agent can relate to a group agent of which she is a member. One of them

---

<sup>31</sup> Nations sometimes carry out these projects not just to secure benefits external to the projects themselves, but because "they" want to be the one to achieve something significant. Think, for instance, of the way John F. Kennedy introduced his plan to travel the moon. Part of his case had to do with benefits that might result from the moon expedition. But much of his pitch centered on the idea that putting a man on the moon would be a remarkable collective achievement, and that the American people should rise to the collective challenge.

is familiar to us from the social contract tradition. The other is less familiar, but I will suggest that it deserves more attention than it has received.

To begin, let us return briefly to non-human animals. If an individual ant and an entire colony are both subjects of natural-historical judgments, questions arise about the relationship between the judgments we make about each. While it is conceivable that our natural-historical judgments would be such that a particular movement on the part of an ant would count as functioning sub-optimally for that individual ant, but not for the colony, we usually judge of animals that live in colonies, packs, and herds, that their characteristic way of functioning *coincides* with the group's. As such, when a particular ant moves dirt in such a way as to maintain a tunnel, we judge that both it and the colony function properly.

The harmony that exists between the individual and the group requires explanation, and there are two kinds of explanation that we can give. The first involves connecting natural-historical judgments about individual ants, the content of which makes no reference to the concept of a colony, to participation in the life of a colony. For instance, imagine that we judge that, "The ant does what it can to ensure that it can dig around in the dirt easily and safely," and we then determine that, for whatever reason—perhaps because the protection that the colony provides allows the ant to concentrate on its digging—an individual ant's ability to dig around in the dirt with ease is much enhanced when she participates in the life of the colony. The fact that dirt-digging is facilitated in a colony would then allow us to connect our natural-historical judgment about the individual with the idea of an ant's acting as a colony member. I call this type of membership "accidental membership."<sup>32</sup>

---

<sup>32</sup> I choose the term "accidental membership" because when an individual is a member of this sort, we endorse a set of natural-historical judgments which leaves open the question whether participating in the colony is part of its functioning well. As an empirical matter, the world turned out in such a way that an ant's ability to dig in the dirt is facilitated by participating in the colony. But it is not essential to the creature's being an ant that it participate in the life of the colony, and in other possible worlds, perhaps, a well-functioning ant can avoid participation in a colony altogether.

Turning now to humans, what would it mean for a rational human being to be “accidentally” a member of society, understood as a group agent? Since a rational human acts for reasons, I suggest that it would mean that the individual has reasons to treat herself as a constituent or to consent to being a member of a group agent, but that the content of the reasons can be articulated without making reference to the group. So we can specify the reason without mentioning the group, but then demonstrate that the reason is best acted upon by joining the group. This is a familiar idea, as it is the basic thesis of social contract theory.<sup>33</sup> Social contract theorists argue that individuals have reason to concern themselves with their own security, natural rights, or freedom (among other things). Then they argue that these things are protected, promoted, or realized by joining and participating in a commonwealth. The individual's membership is then explained either by the fact that she *actually* consents to join

---

<sup>33</sup> Many traditional social contract theorists explicitly state that the parties to the contract are agreeing to become part of a group agent. For instance, in the introduction to *Leviathan*, Hobbes writes:

For by art is created that great Leviathan called a commonwealth, or state (in latin Civitas) which is but an artificial man; though of greater stature and strength than the natural, for who protection and defense it was intended; and in which, the sovereignty is an artificial soul, as giving life and motion to the whole body. . . [T]he pacts and covenants, by which the parts of this body politic were first made, set together, and united, resemble that fiat, or the *Let us make man*, pronounced by God in the creation (pp. 9-10).

In the *Second Treatise*, John Locke writes:

When any number of Men have so consented to make one Community or Government, they are already presently incorporated, and make one Body Politick, wherein the Majority have the Right to act and conclude the rest. For when any number of Men have, by the consent of every individual, made a Community, they have already made that Community one Body, with a Power to act as one Body. (*Two Treatises of Government*, ed. Peter Lasslet (Cambridge: Cambridge University Press, 1988), p. 331.)

Finally, In *The Social Contract*, Jean-Jacques Rousseau writes:

Each of us puts his person and his full power in common under the supreme direction of the general will; and in a body we receive each member as an indivisible part of the whole. At once, in the place of the private person of each contracting party, this act of association produces a moral and collective body made up of as many members as the assembly has voices, and which receives by this same act its unity, its common self, its life, and its will." (*The Social Contract and Other Later Political Writings*. ed., trans. V. Gourevitch (Cambridge: Cambridge University Press, 1997), p. 50, italics removed.)

on the basis of these reasons, or by the fact that she *would* give her consent under ideal conditions.<sup>34</sup> Either way, social contract theory's thesis is that by joining the group agent, an individual protects, promotes, or realizes something she had reason to. This fact, then, allows us to connect our judgment about how the well-functioning individual acts with the idea of her acting as a member of the society.<sup>35</sup> My purpose here is not to render a verdict on the success or failure of social contract theory, as its strengths and weaknesses have been much discussed by others. What I want to suggest instead is that while this is the most familiar way to conceive of the individual's normative relationship to her commonwealth (if we are treating the commonwealth as a group agent), there is another possibility, which is less often discussed.

To illustrate this possibility, let us once again return to non-human animals. Instead of explaining the harmony between our natural-historical judgments about individual ants and our judgments about the entire colony by demonstrating that ants are accidentally members, we might maintain that the content of some natural-historical judgments about individual ants makes explicit mention of the colony. For instance, we might judge that, "When it resides within a particular colony's hill, the ant participates in the life of the colony," or, "The ant, when near a colony that will accept it, does what is necessary to make the colony function properly."<sup>36</sup> Such judgments bring what it is to function properly as an individual ant and what it is to function properly as a colony into immediate alignment. The idea is that once we fully specify what it is to

---

<sup>34</sup> Locke thinks that only those who have given their actual consent are genuine members of the political body. It is a matter of dispute to what extent, if any, Hobbes and Rousseau require actual consent. Kant thinks that the fact that one would consent under certain ideal conditions is sufficient for membership.

<sup>35</sup> It is possible to read Rousseau in a way that makes him sound as if he does not think individuals are accidentally members of his ideal commonwealth, in the sense I have specified here. For Rousseau, freedom is the reason why individuals enter a commonwealth. But there is textual evidence that Rousseau thinks that the commonwealth does not merely promote freedom, but that certain social arrangements are *constitutive* of freedom. If this is right, then to say that an individual has reason to secure her own freedom requires appealing to the notion of society. Importantly, though, the more that commentators emphasize this aspect of Rousseau's view, the more they tend to downplay the role of a social contract in the Rousseauvian account.

<sup>36</sup> I include the caveats about residing within a hill or being near a colony that will accept the ant in order to avoid taking up the issue about an individual ant that is unavoidably separated from any colony.

function properly as an ant, there can be no additional question about whether a particular ant functions well when it participates in the life of the colony. I call this type of membership "essential membership."<sup>37</sup>

It is, I think, most natural to understand the relationship between individual animals and their colonies, packs, herds, and so on, as one of essential membership. This is what Philippa Foot seems to think:

Take, for instance, the dance of the honey bee which tells other bees of a source of food. No doubt an individual bee that does not dance does not suffer from its delinquency, but *ipso facto* because it does not dance, there is something wrong with it because of the part that dancing plays in the life of the species of bee. Similarly, cooperation is something on which good hangs in the life of a wolf, and the free-riding wolf is not behaving as it should.<sup>38</sup>

Natural-historical judgments about individual bees and individual wolves are inextricably bound up with judgments about the hive and the pack. The notion of a well-functioning individual is tied conceptually to the notion of a well-functioning group.

Now let us return, once again, to humans and human society. To treat an individual human as essentially a member of a society, we must believe that the notion of an individual human, acting for the reasons she has, is inextricably connected with the notion of a society. Such an idea is much less familiar to us than is the idea of accidental membership, but I believe that a form of this idea is present in Aristotle's political philosophy. In Book I of the *Politics*, Aristotle writes, "the state is by nature clearly prior to the family and to the individual, since the

---

<sup>37</sup> The term "essential membership" signals simply that it is essential to the creature's being an ant that her functioning as she ought requires cooperating with the colony.

<sup>38</sup> Foot, *Natural Goodness*, p. 35.

whole is of necessity prior to the part."<sup>39</sup> He then proceeds to liken an individual human being to an organ of a living body, whose essence consists in its being part of the whole that it helps constitute.<sup>40</sup> Just as a hand is, Aristotle thinks, defined by its role in the life of an organism, a human is (at least partially) defined in terms of her role in the life of the state. A good human—one who functions as she ought—does not merely manage her own affairs well, but performs well in her role as citizen. So for Aristotle, it isn't as if we first specify what the good life consists in independently of the idea of the state, and *then* explain how the state helps to promote the good life. Rather, the notion of a good human life is tied conceptually to the notion of a political community, society, or state.

The analogy between a citizen and an organ of a living being, along with the idea of a state being prior to the individual, will likely serve mostly to drive away contemporary readers. But in what follows, I will argue for a version of Aristotle's idea (though, admittedly, the argument does not really resemble at all the one that Aristotle eventually advances over the course of the *Politics*).

## 5. Communitarianism

While few contemporary philosophers take seriously the idea of individuals being essentially members of their society, communitarians are arguably an exception. So before proceeding I wish to say something brief about why I believe that their accounts are not fully satisfying, and about what would need to be said in order to give a more persuasive political account, using the Aristotelian point that I mention above.

Michael Sandel, explicitly drawing upon Aristotle, writes:

Some of the special responsibilities that flow from the particular communities I inhabit I may owe to fellow members, such as obligations

---

<sup>39</sup> Aristotle, *Politics*, trans. B. Jowett, in *The Complete Works of Aristotle Volume Two*, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984), p. 1988 (1253<sup>a</sup> 19-20).

<sup>40</sup> Ibid., 1253<sup>a</sup> 20-25. See also *The Movement of Animals*, 703<sup>a</sup> 25-35.

of solidarity. Others I may owe to members of those communities with which my own community has some morally relevant history, such as the morally burdened relations of Germans to Jews, of American whites to American blacks, or of England and France to their former colonies. Whether they look inward or outward, obligations of membership presuppose that we are capable of moral ties antecedent to choice. To the extent that we are, the meaning of our membership resists redescription in contractarian terms.<sup>41</sup>

So Sandel's position is that regardless of whether the individual ever chose to join, she can *just be* a member of a community, and that membership *just does* impact the individual's normative situation. The mere fact that one is born a citizen of Athens, Britain, or the United States can place one under normative pressure to cooperate in the society's pursuits or in discharging the community's obligations, and so it isn't possible to grasp what it is for the individual to act as she ought without grasping that she is a member of one of these societies. Sandel argues for this view by claiming that some of our moral experiences, if they are to make sense at all, require us to posit these "obligations of membership." To illustrate, he relays the story of Robert E. Lee—the native Southerner who opposed secession, but who rejected an offer to lead the Union army, choosing instead to lead the Confederate war effort. Lee's devotion to Virginia—the society into which he was born—compelled him to make the choices that he did. Sandel writes:

One can appreciate the poignance of Lee's predicament without necessarily approving of the choice he made. But one cannot make sense of his dilemma as a *moral* dilemma without acknowledging that the call to stand with his people, even to lead them in a cause he opposed, was a claim of moral and not merely sentimental import, capable at least of

---

<sup>41</sup> Michael Sandel, *Democracy's Discontent: America in Search of a Public Philosophy* (Cambridge, Mass.: Harvard University Press, 1996), p. 15.



weighing in the balance against other duties and obligations. . . A merely psychological reading of Lee's predicament misses the fact we not only sympathize with people such as Lee but often admire them, not necessarily for the choices they make but for the quality of character their deliberation reflects.<sup>42</sup>

The targets of Sandel's argument are liberals who, he believes, deny the possibility of such obligations of membership.<sup>43</sup> But his argument presupposes that Lee indeed faced a genuine moral dilemma, and so presupposes the normative relevance of community membership—the very thing that Sandel is attempting to establish. What we need in order to make the account more compelling is an explanation of why a person is inextricably linked to her community such that her membership has the normative implications that Sandel thinks it does. We need an explanation of why a person who forswears her membership in the community or denies that her community normatively affects her in these ways is making a mistake.

A similar communitarian account is offered by Alasdair MacIntyre. MacIntyre writes, "I will obliterate and lose a central dimension of the moral life if I do not understand the enacted narrative of my own individual life as embedded in the history of my country."<sup>44</sup> So, like Sandel, MacIntyre believes that an individual's self-understanding and identity are naturally and inextricably linked with her membership in a particular society. In fact, he goes on to suggest that all of the moral requirements to which an individual is subject are derived from community membership, and thus that the possibility of being a moral agent *at all* requires one to understand oneself as a member of a community. This is because, MacIntyre claims, any moral

---

<sup>42</sup> Ibid., pp. 15-16.

<sup>43</sup> Sandel is mistaken about what liberalism is committed to, however. Liberals do not deny the possibility of obligations of membership. As I understand the liberal view, one of the reasons that liberals believe freedom of choice ought to be protected is so that individuals will possess the liberty to act from the obligations that they believe bind them as members of particular groups to which they belong.

<sup>44</sup> "Is Patriotism a Virtue?", in *Debates in Contemporary Political Philosophy: An Anthology*, eds. D. Matrauers and J. Pike (New York: Routledge, 2003), pp. 286-300 (quotation from p. 297).

standard, when subjected to rational scrutiny, can be justified only by appealing to the goods of one's particular society. This gives each individual a special stake in the success or failure of her own particular community, nation, or society, and reason to care about whether it does as it ought. The crucial link in this account is the claim that moral standards can only be justified by appealing to the goods of one's particular society. It is for this reason that an individual cannot simply renounce her membership, and claim that she has no role in fulfilling its duties and discharging its responsibilities. But this claim is also the most tenuous link in MacIntyre's account. While he gives some reasons for thinking that the claim is not obviously false, he does not go very far in making the case that it must be true. In fairness to MacIntyre, it isn't clear that he is even trying to provide such an argument, so much as he is trying to lay out a type of view and distinguish it from liberalism.<sup>45</sup> But without more details, we are still left to wonder what goes wrong with an individual who forswears membership, or who denies that the standards of her community have normative purchase on her.

I do not claim to be offering a complete argument against either Sandel or MacIntyre's positions in particular, or against the communitarian position generally. Their views consist of much more than the points I have canvassed here, and there are communitarians besides Sandel and MacIntyre. My purpose is simply to explain why I think that more needs to be said if we are to make use of this particular Aristotelian point in a way that is compelling.

## **6. Society Can Determine What Members Do**

In what follows, I will give an account of the normative connection between society and its individual members, arguing that what a society does (or doesn't do) partially determines what a member does. To give a hint of what I mean by this, think again of the ant moving some dirt. An ant, we said earlier, is essentially a member of its colony. So when we observe it moving some

---

<sup>45</sup> At the beginning of the paper, at least, he seems to suggest he is mostly trying to distinguish two positions, though it is clear which one he favors, and he does eventually offer some arguments for it.

dirt, and try to specify what it is doing, we miss something—we fail to fully specify what it is up to—if we ignore what role the dirt-moving plays (if any) in the life of the colony. The colony thus helps determine that in moving the dirt, the ant is engaging in maintenance of the ant hill. I will argue that we must similarly see an individual human's actions through the lens of her membership in the society (or societies) within which she resides. So when we judge that a human society is  $\phi$ -ing, it is sometimes proper to judge that the members, by virtue of their membership, are  $\phi$ -ing, or that they are participating in  $\phi$ -ing. Because  $\phi$ -ing, here, is understood to be a group action, it is impossible to specify *which* actions are attributable to an individual without making reference to what her society is or isn't doing. So when we say that an individual is under normative pressure to do things that are good or worthwhile, and to avoid doing things that are bad or worthless, we are (I argue) quietly making reference to the individual's society. That is, if an individual is under normative pressure to engage in good action, and what an individual does is partially determined by what her society does, then the normative pressure to engage in good action will include normative pressure to ensure, through her participation, that what her society does is good or worthwhile. In short, the individual's relationship to her society is one of *essential* membership.

Before making my argument, I want to address an objection up front. The paradigm case in which an act, specified using a particular description, is attributable to a rational being is one in which the individual chooses the act under that very description. After all, when giving my Aristotelian account in section 3, I emphasized the importance of rational choice and explained that the capacity to choose for reasons impacts significantly how a human lives a life and thus what we say a human *does*. So an individual might choose to walk across the room in order to get a beer, walk across the room in order to take a call, or walk across the room to leave her lover. We attribute each of these things to her because she chose them. However, most individuals lack the power to choose what their society will do. They lack control over which taxes are levied, over which projects are pursued, over which rights are collectively recognized,

and so on. This raises doubts about whether what a society does can be properly attributed to its members as individuals. Adding to the worry, recall that our moral evaluations of an individual human are evaluations of what she, herself, *does*. But it is a philosophical commonplace that what matters morally speaking, or with respect to whether an individual is a good person, is whether the individual willed the right things, had the right motives, had good intentions, and so on. As Kant famously claims, even if the world conspired to thwart an individual's purposes so that her ends were never realized, the good will, "would still shine by itself, as something that has its full worth in itself."<sup>46</sup> But to allow the society's collective actions to determine what actions are attributed to the individual is to allow something other than the individual's own will, motives, intentions, and so on to determine what she does, and thus, to determine how good a life she lives. So one might object that what the whole society does should not affect our judgments about which acts to attribute to individual members.

While the pull of this objection is undeniable, I believe that it ultimately rests on a mistake. What one does and how morally good one is are *not* completely under one's control. This point is frequently made by those who discuss the phenomenon of moral luck. While moral luck takes many forms, the form I wish to highlight here is what Thomas Nagel calls "luck in the way things turn out."<sup>47</sup> The idea is that what happens as a result an individual's acts, or as a result of an individual's refraining from acting, helps to determine what the individual ultimately does—what we attribute *to her*—even when she lacks control over the results. Consider someone who goes out for a night of heavy drinking with the intention of driving herself home afterward. On her return trip, when she swerves off the road, it is a matter of luck or chance whether a pedestrian is standing on the sidewalk where the car swerves.<sup>48</sup> But despite

---

<sup>46</sup> Immanuel Kant, *Groundwork of the Metaphysics of Morals*, ed., trans. Mary Gregor (Cambridge, UK: Cambridge University Press, 1996), p. 8 (4:394).

<sup>47</sup> Thomas Nagel, "Moral Luck," in *Mortal Questions* (Cambridge, UK: Cambridge University Press, 1979), pp. 24-38 (quotation from p. 28).

<sup>48</sup> *Ibid.*, p. 29.

the fact that she made identical choices either way, the presence or absence of a pedestrian will make all of the difference when determining whether the driver was merely “driving under the influence,” or whether she instead committed manslaughter. It will make the difference between our judging that she merely acted recklessly, and our judging that she killed someone.

Cases like this are ubiquitous. Imagine that I push someone lightly with the intention of merely humiliating her, and she trips on a branch as she is stumbling, hits her head, and then dies. The fact that some bad luck was involved might mitigate my responsibility—few would deny *that*. (I won’t, after all, be tried for first degree murder). But we nevertheless would say that I killed her. Her death is my responsibility—a fact I will have to live with for the rest of my life, despite the fact that but for the presence of the branch I would merely have committed a minor assault. Or imagine that I intentionally incite a riot, and the crowd then carries out horrible deeds that I never anticipated. In such a case, we won’t judge that I merely shouted some words. We will also say that I helped to bring about mayhem, death, and destruction, and the victim’s families can reasonably lay these consequences at my feet. In short, what one does, and what one is morally responsible for, are partially determined by things outside one’s control. This is because we do not simply think that a person’s immediate movements—like a wiggle of her toe or a movement of her arm—are attributable to her. We also think that some of the consequences of those movements should be incorporated into the description of what she does.<sup>49</sup> And sometimes an individual lacks control over those consequences. So, as Nagel concludes, “A person can be morally responsible only for what he does; but what he does results

---

<sup>49</sup> At one point, Hegel remarks that when we act, “[the consequences], as the [outward] shape whose soul is the end to which the action is directed, belong to the action as an integral part of it.” (Hegel, *Elements of the Philosophy of Right*, p. 145). So while we can isolate certain momentary acts undertaken by an agent, the true essence of what is being done in those isolated moments often depends upon consequences—states of affairs that obtain at a time or place remote from the one in which the agent does her part. Hegel says that arson provides a clear illustration (*ibid.*, pp. 147-148). If one knows that an individual sets a piece of wood on fire, but fails to grasp the eventual consequences, then one misses out on the fact that this very act of setting a piece of wood on fire *is an arson*.

from a great deal of what he does not do."<sup>50</sup> Of course, one *could* argue that consequences that result from luck *should not* be attributed to the agent, insisting that only what she has directly chosen should be. But Nagel's point, I believe, is that this would require enormous revision to the way we think about rational agency. It would require that the drunken driver say, "I didn't kill that pedestrian. I merely drove recklessly," or that the person who shoved the other say, "She died, but I didn't do it. I merely gave her a push," or that the person who incites the riot say, "Mayhem ensued, but I didn't bring it about. I merely said some words." All of this is quite implausible. Something in the idea of rational agency requires that an individual identify more closely with the consequences of her chosen acts than this.

So it is *not* true that what a person does is wholly determined by the choices she makes. Events external to an individual's choice can "infect" what she does. This, at the very least, opens up the possibility that what an individual's society does also "infects" what she does—that what she does results from a great deal of what her society does. If this is right, then what an individual's society does will have a direct bearing on whether what an individual is doing is good or worthwhile.

One way to argue that the acts and omissions of a whole society can determine, at least in part, what its individual members do would be to give a general principle stating conditions under which an event is attributable to an individual, and then show that the conditions are met when a society acts. Unfortunately, as far as I am aware, no one has proposed a satisfying principle of this sort. So rather than taking this approach, I will instead proceed as we did with moral luck—examining specific examples and demonstrating the implausibility of denying that what the society does "infects" what the member does. I believe that my case is most clearly made by thinking about economics, and so in what follows, I will offer two examples pertaining to an individual's economic activity.

---

<sup>50</sup> Nagel, "Moral Luck," p. 36.

The first example pertains to an individual's participation in the economic practices that her society puts in place and sustains. Such practices determine which objects individuals may permissibly use and exclude others from using—they determine who owns what within a particular society. Examples of possible practices include allowing the individual to exclude others from land with which she mixes her labor, allowing the individual to keep objects she acquires through an uncoerced agreement with another party, or allowing the individual to control whatever amount of money she makes after she pays her taxes. Any society will have practices of this sort, even if they are largely informal.

Particular sets of practices often produce predictable consequences. For instance, a particular kind of economic system might, by imposing few constraints on the kinds of exchange that individuals can engage in and the amount of property they can accumulate, predictably create a very rich, Aristocratic class, leaving some people poor. Or, a particular kind of economic system, by investing a single person or institution with too much power to determine who owns what, might predictably lead to scarcity of needed goods and services. Because the society has the capacity to sustain practices intentionally, the consequences of its practices may plausibly be attributed to *it*. It is less obvious, however, whether the consequences “infect” the individual members' action—whether they help determine what the individual members do. After all, although the individual may be causally related in some way to the consequences, it is not as if every consequence of an individual's action is attributable to her—some of my actions initiate causal chains that produce results so remote that no one would attribute them to me. So we might wonder: does buying an item, collecting a paycheck, or investing in a mutual fund really implicate an individual in the consequences brought about by her society's economic practices?

To answer this question, I will make two brief points to help characterize the relationship between an individual and the consequences brought about by her society's economic practices. The first is that many of an individual's economic actions are defined by her society's practices themselves. To fully understand what I am doing when I purchase an object or some land, one

must have some understanding of my society's economic practices—practices that specify who has the right to exclude others from what, what counts as a legitimate transfer, and so on. So to buy and sell objects within one's society, or to sell one's labor or purchase services, involves intentionally partaking in a practice. The second point is that these practices are upheld and sustained through members' participation. When we say that there exists an economic practice within a particular society, we mean that many or most of the members participate. So if no one recognized others' right to exclude people from using objects that they purchased, or no one recognized that part of a wage-earners paycheck (the part taken out for taxes) is rightfully controlled by the government, then these practices specifying who owns what would simply cease to exist.

So the question, in the end, is not whether a causally remote result can “infect” the action of an individual who engages in economic activity by virtue of the fact that her action happens to be the upstream cause of the result. Rather, the question is whether a result can “infect” an individual's action by virtue of the fact that the result is brought about by a practice sustained through the individual's intentional participation. As I have said, I do not have a principle to apply to answer this question. Nevertheless, I believe that our concept of rational agency does allow such consequences to “infect” an individual's actions. To illustrate, consider arguments made by defenders and detractors of particular economic schemes, such as capitalism. Defenders often argue that by participating in the free market, an individual helps to benefit others and that they need not worry that their activity is “infected” by harms or injustices generated by the market. Hegel's admiration for capitalism, for instance, is based, in part, on his belief that within the market, members of civil society can carry on knowing that they are advancing the interests of their fellow citizens. To give another example, Rawls thinks that one of the virtues of a market designed to adhere to the difference principle is that each market actor is free to pursue her own good, confident that in so doing she is helping to advance the interests of the least advantaged, thus giving justice to her fellow citizens. On the other hand, those who



criticize capitalism sometimes emphasize that those who benefit have “dirty hands,” so to speak. The market, it is argued, is unjust and inhumane, and as a result, individuals’ economic activity is tainted. All of this seems to rely upon the tacit assumption that an individual rational agent cannot simply say, “Since I lack control over which economic practices my society adopts, and because my actions are not directly responsible for the consequences, my actions are not implicated in the consequences brought about by my society’s economic practices.” That is, we seem to think that a rational agent’s actions—what it is that she does—is partially determined by her society and the practices that it adopts and sustains.

Consider, as a second example, the fact that society has the capacity to collect taxes—to take a portion of the wealth generated from its members’ economic activity—and allocate them toward projects that no individual alone is in a position to carry out. When it exercises this capacity, we often say of a society that it  $\phi$ ’s through the use of taxpayer dollars. When the individual engages in the activity that helps generate the wealth that is subsequently allocated to some project, what then becomes attributable to her?<sup>51</sup> The building of roads, the maintenance of schools, and creation of an army all result more or less directly from the taxes that she and others pay. Indeed, taxes are usually levied and turned over for these specific purposes—they are the reason why the taxes are collected to begin with. Given all of this, can our taxpayer plausibly deny that the projects to which her taxes are allocated “infect” what she does?

Consider, as an illustration of how we typically think that a rational agent relates to taxpayer funded projects, debates about the use of taxes for controversial purposes—for instance, to pay for abortions. Many who seek to outlaw all abortions make a more urgent complaint about taxes being used to pay for the procedure. The complaint seems not merely to

---

<sup>51</sup> Admittedly, there are fraught issues here about *who* generated the wealth that is taxed. Certainly the individual who earns an income and pays the tax plays a role. But I am inclined to say that all of society’s participants help generate the wealth that is eventually taxed. This is, after all, why many philosophers writing about social justice believe that we need principles of justice to determine how wealth is distributed. Because the wealth is a social primary good (to use Rawls’s terminology) created through collective interaction and cooperation, “we” need to determine how to distribute it.

be that the money could be put to better use, but that by funding abortions, society's members actually become implicated in providing abortions. This is, I believe, the reason why many who are not usually willing to accommodate the anti-abortion agenda are more likely to accommodate it in this instance.<sup>52</sup> What everyone intuitively seems to be picking up on is that taxes are a way of extracting a direct contribution for a project from individuals whose activity sustains the economy. Granted, individuals are usually coerced into paying taxes, and will face fines or imprisonment for not doing so. But this does not undermine the present point. As I understand them, abortion opponents object to being put in the position of having to choose between involving themselves in the abortion-providing business and being punished, or to being put in the position of having to choose between engaging in commerce and helping to provide abortions. After all, to say that an individual was coerced into  $\phi$ -ing does not imply that she didn't  $\phi$ . Indeed, the coerced individual's complaint rests on the assumption that when she is coerced into  $\phi$ -ing, she ultimately  $\phi$ 's. Rarely is it argued that because the government is the entity that ultimately allocates the funds for the abortions, the taxpayer or citizen should not be concerned that *she* is implicated. This is, I think, because we generally agree that what a society does with the tax dollars it collects helps determine what the individual taxpayers do.<sup>53</sup> (And obviously, here, abortion is simply an example. I could have made the same point using the example of funding a war that some believe is unjust.)

While I have only examined two examples, I think a picture is beginning to emerge. Just as the drunken driver, as a rational agent, cannot dissociate herself from the death of the person standing on the sidewalk, an individual living in society cannot dissociate herself from the

---

<sup>52</sup> For instance, the Hyde Amendment has, since 1977, proscribed federal funding of abortions in the United States, and has been consistently supported by many legislators who otherwise support a woman's right to choose.

<sup>53</sup> To be clear, one could agree with the claim that an individual who pays her taxes involves herself in the projects that those taxes fund, while holding that taxes should still fund projects and services that some individuals object to on moral grounds. So it isn't that one must, upon endorsing my view about attributability, also endorse the Hyde Amendment.

consequences of the economic practices that her society institutes and sustains, or from the endeavors funded through tax dollars. While it is difficult to give a precise theory about why, exactly, this is so, something about our conception of rational agency seems to require acknowledging that what a society does helps to determine what its members, as individuals, do.

Now, at last, I will use the argument made above to explicate an individual's normative relationship to her society, and to argue that she is essentially a member. Rationality, we said earlier, demands that an individual take the fact that  $\phi$ -ing would be good or worthwhile to place her under normative pressure to  $\phi$ , and the fact that  $\phi$ -ing would be bad as a reason to refrain from  $\phi$ -ing. As I argued above, it is often not possible to specify fully what an individual is doing, or failing to do, without referring to what her society is doing. To say of an individual that she ought to ensure that the action she engages in is good or worthwhile, thus, is already to claim that the well-functioning individual does what she can to make society function as it ought.

Let me say a bit more to help flesh this out. Functioning as a rational agent requires that an individual aim to ensure that her own actions are good or worthwhile. Functioning as a rational being does not, however, require that an individual aim to ensure that everything else in the world functions as it ought. A rational human is under no normative pressure to, for instance, ensure that a particular knife or car functions as it supposed to, or to make sure that a plant or a bird function the way they are supposed to. If one encounters a malfunctioning artifact or organism, one can sensibly ask, "Why should I concern myself one way or another with how it functions?" But if an individual's society is functioning poorly, and failing to act in ways that are good or worthwhile, I suggest that we now have a ready answer to the question, "Why should I concern myself one way or another with how it functions?" The answer is that so long as we admit that the individual is under normative pressure to act in ways that are good or worthwhile, she ought to concern herself with what her society does, because her society

partially determines what she does. So, when the individual's society does something that is good or worthwhile, the individual is under normative pressure to cooperate—to pay the relevant taxes, to obey the laws, and so on. When it refrains from doing something that is worthwhile, she is under normative pressure to do things in her power to get it to act in ways that are worthwhile. When it acts in ways that are positively bad, she is under normative pressure to do what she can to get it to give up on the bad action. These are all things that a well-functioning human will do. Just as an ant that does as it ought is one that helps to make the colony function as it ought, an individual human that does as she ought is one that helps her society function as it ought.

### 7. The Commonwealth as a Natural Agent

I will conclude, here, by pulling together various strands of the argument that I have made in this chapter, and then I will spell out the account's implications for how we ought to understand the state and its proper role. It is perfectly sensible to ascribe agency to groups, and to do so regardless of whether the group's members ever decided to create an agent. Doing so does not require appealing to anything supernatural. It simply requires treating the group to be a bearer of a particular form, with a characteristic way of functioning.

An individual and her society are not two entirely distinct entities, whose forms can be grasped independently of one another. The individual is *essentially* a member of the society to which she belongs, in the sense that when we specify what it is for her to function properly, we must mention actions taken (or not taken) as a member of her society. Her own activity, that is, must be seen through the lens of her membership in the society. Because of this, she is under normative pressure to participate in (and to cooperate with) action that it would be good for the society to carry out, and to do what she can to ensure that her society engages in good or worthwhile action. So our judgments about what it is for her to function well are inextricably bound up with our judgments about what it is for the society to function well.

Acting together as a society will often require making collective decisions in a formal way, collecting taxes, promulgating rules and regulations, solving collective action problems, and enforcing laws. Doing these things requires a state, with institutions capable of structuring deliberation and enforcing decisions. The society, in short, needs institutions that will turn it into a fully-fledged political community—it needs institutions that will turn it into a commonwealth. Earlier, I argued that a society, as a group agent, is under normative pressure to do things that are, quite generally, worthwhile. This might include doing things that a minimal state would do, such as protecting individual rights. But it might also include more ambitious pursuits such as supporting the arts, curing cancer, achieving artistic excellence, and so on. Insofar as these things require use of the state, normative pressure to do these things *just is* normative pressure to use the state to do them. My view is thus that the role of the state is potentially quite expansive.

The account that I have given here, however, leaves two major issues in need of resolution. The first is that even if one accepts, as a general matter, the argument I have made, one might still maintain that additional considerations limit dramatically which worthwhile projects the state may permissibly pursue. For instance, one might argue that we wrong the taxpayer by pursuing a project for no other reason than that it is good or worthwhile, or that individual members have the right to choose whether or not to engage in such collective endeavors. So whatever normative pressure is generated from one's being essentially a member of a society might be overridden or trumped by these other considerations. The second issue is that even if one accepts that individuals are under some level of normative pressure to engage in worthwhile collective action, there will be other normative pressures to which the individual is subject as well. An individual might, after all, recognize the importance of some collective endeavor, while valuing much more highly her own individual projects. But if, in the end, each individual is permitted to refrain from participating in the commonwealth's collective endeavors, then the implications of my argument here turn out to be rather limited. So

something needs to be said about how (or whether) an individual who values her own individual projects more than her society's collective endeavors comes under normative pressure to prioritize the collective projects over her own. That is, we need an account of why an individual ought to cooperate with the commonwealth when it makes a decision about what to do, even when she would rather carry on with her own individual projects. Much of the rest of the dissertation will be spent exploring these issues.

## Chapter 3: Duties to the Self

### 1. Introduction

Those who claim that the commonwealth is an agent rarely offer up the individual agent as a model for understanding how an entire political body should act generally. In fact, philosophers typically warn against applying too generously what we know about individual agency and individual practical reasoning to entire political communities. For instance, it is generally agreed that an individual agent may permissibly balance burdens and benefits within her life in order to advance her overall good. But this should not, we are warned, cause us to think that a commonwealth may similarly balance burdens and benefits among its members in order to achieve *its* overall good.<sup>1</sup> It is also generally agreed that an individual may permissibly commit herself to a particular course of action in the pursuit of ends that she chooses. But again, this should not, we are warned, cause us to think that the commonwealth may permissibly commit itself to a particular course of action in the pursuit of ends that *it* chooses.<sup>2</sup>

There are, no doubt, many reasons that one might think that the individual and the commonwealth differ in these fundamental ways. But I suspect a common one is the belief that while an individual person cannot wrong herself (and is thus free to pursue her own good and her own ends without risk of wronging herself), a commonwealth is always in danger of wronging its members as it pursues the common good or collective ends. That is, by imposing upon its members the burdens necessary to achieve the common good, and by enlisting members in the pursuit of ends that they didn't themselves choose, the commonwealth risks wronging the individuals that constitute it. So despite my argument in the previous chapter that

---

<sup>1</sup> John Rawls, *A Theory of Justice, Revised Edition* (Cambridge, Mass.: Harvard University Press, 1999), pp. 19-24; Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 32-33.

<sup>2</sup> Kant argues that an individual agent has an innate right to pursue her own ends. But while Kant admits that an entire people possesses a will, according to Arthur Ripstein at least, he denies that a people has the right to choose ends for the whole. (See Arthur Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy* (Cambridge, Mass.: Harvard University Press, 2009), p. 196.)

both the individual and the society as a whole are under normative pressure to live out a life of activity that is good or worthwhile, one might think that the similarity between the two disappears once we notice that the commonwealth cannot, without committing a moral wrong, do many of the kinds of things that a life of worthwhile activity typically consists in.

With an eye to demonstrating that the individual agent is a better model for understanding the commonwealth than is commonly supposed, I will, in this chapter, argue that an individual *can* wrong herself, and thus that she is morally constrained in ways that the commonwealth is. If true, this is significant. Moral duties are thought by many to impose particularly weighty, stringent, or even overriding requirements to act or to refrain from acting in particular ways—requirements that many think explain why a commonwealth may not treat its members in certain ways, even to promote important goods or to achieve worthwhile ends. If the individual, despite having moral duties to herself, may permissibly balance burdens and benefits in order to pursue her overall good, then it raises the possibility that the commonwealth may also balance burdens and benefits in pursuit of society’s overall good. And if the individual, despite having moral duties to herself, may permissibly commit herself to pursuing ends of her choosing, then it raises the possibility that a commonwealth may commit itself to pursuing its collective ends as well. In fact, in subsequent chapters I will argue that commonwealths can indeed acquire these permissions, and I will look to the individual agent as a model to help explain how they can come to acquire them.

Now, the topic of duties to the self isn’t widely discussed these days. One reason, I suspect, is that many philosophers think it obvious that such duties don’t exist.<sup>3</sup> While everyone admits that an individual has self-regarding reasons—for instance, reasons to promote her own

---

<sup>3</sup> Take, for instance, Bernard Williams. He writes, “Morality encourages the idea, *only an obligation can beat an obligation*.” This, Williams complains, leads to morality’s crowding out important endeavors normally associated with a good life. He suggests that the only way for the moralist to make room for a person to have something that resembles a normal good human life is to introduce, “one of those fraudulent items, a duty to myself.” The fact that such an item is needed is, Williams seems to think, a *reductio ad absurdum* of the moralist’s position. (See his *Ethics and the Limits of Philosophy* (Cambridge, Mass.: Harvard University Press, 1985), pp. 200–202.)



welfare or to treat herself with dignity—many deny that these are "moral reasons," or "reasons of duty," as opposed to "reasons of prudence." In this chapter, I challenge the prevailing view, arguing that there are indeed moral duties to the self, and that with them come the weighty, stringent, or overriding requirements to act (or not to act) mentioned earlier. I begin in section 2 by discussing a feature of moral reasons that many philosophers think distinguishes them (or, at least, an especially important subset of them) from other practical reasons, which is their second-personal character. In sections 3 and 4, I argue that the notion of a second-personal reason that is self-regarding is coherent, and that such reasons exist. Finally, in section 5, I tie a moral reason's stringency to its second-personal character, and then argue that the individual and the commonwealth face a similar predicament when deciding what to do.

## **2. Morality's Second-Personal, Bipolar Character**

Let me begin by laying out two examples of what I take to be paradigmatic kinds of moral wrongs. First, there are cases in which an individual disregards the interests of others while seeking to advance her own interests, causing harm to other individuals in the process. For instance, imagine I am a pharmacist who knows that the community needs a particular drug in order to prevent mass suffering, and I deliberately create a shortage in order to marginally increase my profits. If all I have to say in my defense is that my interests count for everything, and everyone else's count for nothing, then the suffering members of the community will be justified in feeling morally wronged. Second, there are cases in which an individual places no deliberative importance on the fact that another has ends or goals, and thus undermines the other person's ability to exercise her autonomy. For instance, imagine that because I aim to maintain a Christian community, I use brute force to prevent you from building a mosque. If all I have to say in my defense is that my ends counts for everything, and your autonomy counts for nothing, then you will be justified in feeling morally wronged.

Now consider two very similar cases, that differ only in that they involve one person rather than multiple people. First, imagine an eighteen-year-old who smokes habitually because of the pleasure it affords her in the present. She believes she will develop emphysema later, and suffer greatly as a result. But she disregards this entirely because the benefits are immediate and the burdens are distant. Second, imagine a variation of a case considered by Derek Parfit, in which a socialist Russian nobleman suspects he will undergo a sincere shift in his values as he ages.<sup>4</sup> Discounting entirely the more conservative ends he expects to have in the future, as well as his ability to autonomously pursue them in his old age, he commits his fortune in such a way as to limit his control over it as he ages—to make it so that he will not be able to use it to further his conservative goals.

These two cases parallel closely our paradigmatic cases of moral wrongdoing, in that the individuals involved arguably err in their practical reasoning in similar ways—ignoring considerations pertaining to welfare and autonomy that should, at the very least, factor into their deliberations. Many would insist, though, that the mistakes made by the smoker and the nobleman are mere *prudential* mistakes, rather than *moral* mistakes.<sup>5</sup> My contention is that both sets of cases involve moral mistakes. The reasons an individual has to promote her own welfare and to protect her own autonomy are not merely prudential, but stem from moral duties that she has to herself. But what differentiates moral considerations from non-moral ones? There exists little agreement on how to answer this question (or even on whether there is an answer). However, many philosophers have suggested that moral reasons—or at least an interesting and important subset of them—are distinguished by their "second-personal" or

---

<sup>4</sup> Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), pp. 327-328. Importantly, it is not that the nobleman thinks he will eventually give into temptation and betray his values. Rather, he thinks he will experience a genuine shift in his values. So he is trying to ensure that his future self will not be able to act on the values he will genuinely endorse at the time.

<sup>5</sup> For a famous example, see Chapter 4 of John Stuart Mill, *On Liberty*, in *On Liberty and Other Writings*, ed. Stefan Collini (Cambridge, UK: Cambridge University Press, 1989), esp. pp. 78-79.

"bipolar" character.<sup>6</sup> When I claim that there are duties to the self, what I mean is that individuals have self-regarding reasons that are second-personal in this sense. Before proceeding, then, I must say a bit about what I mean by this.

Michael Thompson begins his discussion of this topic by noticing that there is something common to judgments of the form "*X wrongs Y*," "*X has a right against Y*," and, "*X has a duty to Y*."<sup>7</sup> Such judgments express a relation between two individuals that appears to be absent from judgments of the form "*X ought to  $\phi$* " or "*It would be good for X to  $\phi$* ," even when  $\phi$ -ing involves doing something *to Y*, or *for Y*. Those who make this point often maintain that bipolarity is an irreducible notion, making it difficult to define it precisely. But to help illustrate what is distinctive about this species of judgment, consider a case discussed by Stephen Darwall. Imagine an individual who steps on another's foot, causing her pain. The offending individual, most would agree, has reason to remove her foot. An argument that she does could be made by noticing, with G. E. Moore, that, "doing so will cause more good to exist in the Universe than any possible alternative."<sup>8</sup> As such, the owner of the offending foot has a strong reason to remove it—a duty to do so, Moore thinks. The judgment arrived at here is of the form, "*X ought to  $\phi$* ," or "*X has a duty to  $\phi$* ." But according to Darwall, a foot-stomper who acts *merely* on the basis of this type of judgment is insensitive to the fact that the victim herself possesses special authority to demand of the offender that she remove her foot. Darwall thus introduces the notion of a second-personal reason—a kind of reason that, by acting upon it, one acknowledges that the

---

<sup>6</sup> Some recent works that emphasize this aspect of morality include: T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998); Stephen Darwall, *The Second-Person Standpoint: Morality, Respect and Accountability* (Cambridge, Mass.: Harvard University Press, 2006); Michael Thompson, "What is it to Wrong Someone: A Puzzle About Justice," in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, eds. R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (Oxford: Oxford University Press, 2004), pp. 333-384; R. Jay Wallace, "The Deontic Structure of Morality." Manuscript available at: <http://philosophy.berkeley.edu/file/2/Deonticstructure.pdf> (accessed September 1, 2012).

<sup>7</sup> Thompson, "What is it to Wrong Someone?" p. 336.

<sup>8</sup> G. E. Moore, *Principia Ethica* (New York: Dover, 2004), p. 148.

offender owes it *to the victim* to act.<sup>9</sup> The genuinely moral individual, he thinks, is sensitive to second-personal reasons, and thus makes judgments such as: "If I fail to remove my foot, I will *wrong* this person," or, "This person *has a right against me*, requiring me to remove my foot," or "I *owe it to* this person to remove my foot."

To help illustrate the intuitive sense in which second-personal reasons and bipolar judgments differentiate moral considerations from considerations we tend to think are non-moral, consider a pair of examples offered by R. Jay Wallace. In the first, a tired individual is inclined to stay in for the evening. But, on this night only, there is a screening of *Rocco and His Brothers* at an elegant old theater, and never again will she have the opportunity to view the film under such auspicious conditions. Thus, despite being tired, she has decisive reason to go out. In the second, an individual has taken out a loan and has promised to make payments to the lender on the fifth of the month. Today is the fifth. Thus, despite having other things she would like to do with her money, she has decisive reason to make the payment.<sup>10</sup> Wallace lays out these examples in order to highlight that while both the potential movie-goer and the money-borrower are under normative pressure to act, we intuitively recognize a distinctive sort of practical necessity in the loan case that is absent from the movie case—a sort practical necessity characteristic of instances in which one has a *moral duty*. Some caution is required here. It isn't that the person is *irrational* if she fails to repay the loan, but not if she fails to attend the movie-screening. A person can, after all, behave irrationally merely by failing to do what she knows she has decisive reason to do, and thus by foregoing the movie screening. What differentiates the two cases, Wallace suggests, is that in the loan case, "Your reason for payment is not a free-standing normative consideration, but part of a nexus linking you and your creditor." Moral reasons, he proposes, "are constitutively implicated in complexes of relational (or "bipolar")

---

<sup>9</sup> The foot example is discussed in Darwall, *The Second-Person Standpoint*, pp. 5-10.

<sup>10</sup> Wallace, "The Deontic Structure of Morality," pp. 3-4.

normativity.”<sup>11</sup> I do not want to place too much weight on the thesis that *all* reasons that we designate “moral reasons” have this bipolar character, as the word “moral” is used in a variety of ways.<sup>12</sup> But those reasons that do have this character certainly comprise an important subset of the reasons that we typically designate “moral,” and it is these reasons that will primarily interest me here.

### 3. The Possibility of Relating to Oneself Second-Personally

In claiming that there are duties to the self, then, what I mean is that we can properly make judgments like, “X *wrongs* Y,” “X *has a right against* Y,” and, “X *has a duty to* Y,” in cases where X and Y are identical. That is, sometimes when an individual ought to act out of concern for her own welfare or out of self-respect, she *owes it to herself* to do so.

Of course, there is no problem, in general, with suggesting that a person can relate to herself in a way that she typically relates to others. A person can love herself, care for herself, know herself well, stand in judgment of herself, and so on. There is, however, a famous objection to the very possibility of an individual relating to herself morally. Kant (who ultimately argues that there *are* duties to the self) articulates the objection by pointing out that, “the one imposing the obligation could always release the one put under the obligation, so that (if both are one and the same subject) he would not be bound at all to a duty he lays upon himself.”<sup>13</sup> More recently,

---

<sup>11</sup> Ibid.

<sup>12</sup> Even among authors whose work focuses most directly on morality’s second-personal character, there are some who hesitate to claim that *all* moral wrongs must involve these bipolar relations. Scanlon says that the range of criticisms that people typically recognize as “moral” is quite broad, and that his book considers only the more narrow topic of, “the morality of what we owe to each other.” (See his *What We Owe to Each Other*, pp. 6-7.) Thompson similarly writes, “[J]ustice isn’t the *only* virtue, and so I can intelligibly be said to do wrong or go wrong or act wrongly, morally speaking, even when no one is wronged.” (See his “What is it to Wrong Someone?” p. 339.) Darwall and Wallace seem more inclined to insist that if something is a moral reason, it must be second-personal, though the evidence from Darwall’s writings is mixed on this point.

<sup>13</sup> Immanuel Kant, *The Metaphysics of Morals*, ed., trans. Mary Gregor (Cambridge, UK: Cambridge University Press, 1996), p. 173 (6: 417). Kant initially responds to this objection by asserting that all duties depend upon duties to the self. So if there are no duties to the self, then there are, Kant thinks, no duties at all. But since we know that there are duties, we must infer that there are duties to the self. So Kant

Marcus Singer has made a similar argument: "(1) If A has a duty to B, then B has a right against (or with respect to) A; (2) if B has a right against A, he can give it up and release A from the obligation; and (3) no one can release himself from an obligation. From this it [follows] that if one could have a duty to oneself. . . that this alleged duty to oneself would be a 'duty' from which one could release oneself at will, which is self-contradictory."<sup>14</sup> The Kant-Singer objection, then, is that because an individual would be able to release herself from any apparent duty to herself, the duty would necessarily fail to normatively bind her in the way characteristic of moral obligations. Thus, the notion of a duty to the self appears internally incoherent.<sup>15</sup>

The Kant-Singer objection gets purchase because the two poles of the purported bipolar relationship are identical. It is the fact that the potential wrongdoer is identical to the victim that explains why the potential wrongdoer cannot be normatively bound by the purported duty. One way to escape the objection, then, would be to claim that distinct temporal stages within a life are non-identical, metaphysically distinct entities.<sup>16</sup> So, in our examples, the stage of the smoker's life that actually does the smoking would be treated as an entity distinct from the one that suffers from emphysema, and the stage of the nobleman's life that constrains the fortune would be treated as an entity distinct from the one that is eventually constrained. The earlier

---

concludes that the argument purporting to expose the self-contradictory nature of duties to the self is an antinomy. In attempting to resolve the antinomy, Kant distinguishes between the noumenal and phenomenal self. This allows him divide the self into two, with one aspect possessing authority and the other owing (morally) its obedience to the other. Dividing the self in order to explain how an individual can have a duty to herself is a strategy that is attractive to me, but my argument will appeal to temporal divisions within a life rather than to the Kantian notions of phenomena and noumena.

<sup>14</sup> Marcus Singer, "Duties and Duties to Oneself," *Ethics* 73 (1963), pp. 133-142 (quotation from p. 133).

<sup>15</sup> One might have misgivings about the assumption, made by both Kant and Singer, that an individual can always release another from duties owed to her. Some believe, for instance, that persons have inalienable rights, and that A necessarily wrongs B if she violates B's inalienable rights, regardless of whether B consents. While I am sympathetic to this worry, I am not going to base my argument upon it. Whether or not inalienable rights exist is a matter of deep controversy, and even those who think we have such rights believe that they constitute a rather limited subset of all our rights.

<sup>16</sup> Parfit, for instance, denies that there is anything more to a temporally extended person than clusters of mental states at distinct moments, and the causal connections between them. So Parfit could treat a cluster of mental states at one moment as an entity that is distinct from (or that is not identical with) a cluster of mental states that exists later on—even when both belong to what we would normally consider to be the same person. (See *Reasons and Persons*, Part III).

stage owes it to the later stage to act in a particular way, and, since the earlier stage is not identical to that later stage, it cannot release itself from the duty. The objection is escaped, then, by reinterpreting our examples so that when we say, "X *wrongs* Y," we *do not* consider X to be identical with Y. Escaping the Kant-Singer objection in this way, however, would be extremely controversial. The claim that the young smoker and the young nobleman are not identical with their older selves is certainly in tension with our normal way of speaking. And furthermore, many metaphysicians have argued that the idea that a temporally extended object is constituted by metaphysically distinct stages, or temporal parts, is either implausible or incoherent.<sup>17</sup>

As an alternative, I wish to give an explanation of how our smoker and our nobleman, despite retaining their identity from one moment to the next, can relate to themselves in a way that mirrors the way that two distinct individuals relate to one another—in a way that precludes them from releasing themselves from the duty. To do this, I will argue that it isn't actually their metaphysical distinctness that ultimately explains why two distinct persons can relate to each other second-personally. It is, rather, the fact that they each have distinct "perspectives" or occupy distinct "standpoints." Then I will show that an individual person herself has a succession of distinct perspectives over time. This, I will argue, ultimately enables her to relate to herself second-personally. In this way, the Kant-Singer objection will be answered.

In the opening passages of his book, Darwall defines the second-person standpoint as a "perspective" from which one makes or acknowledges claims on the will and conduct of someone who has a different "perspective."<sup>18</sup> Second-personal reasons are generated either when an individual *actually* issues a justifiable demand from this standpoint, or could

---

<sup>17</sup> See Judith Jarvis Thomson, "Parthood and Identity Across Time," *Journal of Philosophy* 80 (1983), pp. 201-220; Peter van Inwagen, "Four-Dimensional Objects," *Nous* 24 (1990), pp. 245-255. For a defense of the claim that temporally extended objects are composed of temporal parts, see Ted Sider, *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford: Oxford University Press, 2002).

<sup>18</sup> Darwall, *The Second-Person Standpoint*, p. 3.

*hypothetically* issue one.<sup>19</sup> To illustrate, let us first revisit the inter-personal case. Imagine that Smith has a mild desire to stand where Jones is, and so steps on Jones's foot, which happens to be sprained. (Let us stipulate that the incident occurs on a crowded bus, where invasions of one's personal space are considered to be expected and permissible so long as no one is caused any pain.) From Smith's perspective, things go well, as his desire is satisfied. From Jones's perspective, things go badly, as having his sprained foot trampled upon is very painful. Jones, as the occupier of the perspective from which things go badly, calls attention to the fact that he is in pain and demands that Smith remove his foot. If Smith refuses, then Jones judges that he is wronged by Smith. If Smith obliges, then Smith takes himself to be acting on a second-personal reason, recognizing that he owes it to Jones to remove his foot. Obviously Smith and Jones are distinct persons. But, if I understand Darwall correctly, it is the fact that the two have distinct perspectives that enables them to stand in this sort relationship to each other. There is the fact that there are distinct perspectives, the fact that the interests had from those perspective conflict, and the fact that on the basis of the interests had from his perspective, Jones makes demands that Smith, from his perspective, take to be reason-giving. This is how a second-personal reason is generated in this case.

It would seem, then, that the possibility of an individual taking up a Darwallian second-personal stance with respect to herself, and of generating second-personal reasons from it, could similarly be explained by her occupying two distinct perspectives from which demands could be issued and acknowledged. To see that this is indeed possible, consider the fact that some

---

<sup>19</sup> Darwall's language is voluntarist throughout much of his book, leading some to think he is overly reliant on the individual making actual demands. (See R. Jay Wallace, "Reasons, Relations, and Commands: Reflections on Darwall," *Ethics* 118 (2007), pp. 24-36). But he does, in some places, appeal to the idea of a hypothetical moral community that makes demands on its members. (For a discussion of this point, see Darwall, "Reply to Korsgaard, Wallace, and Watson," *Ethics* 118 (2007), pp. 52-69, (esp. pp. 64-65)). One needn't, however, endorse Darwall's specific Kantian proposal to accept the idea that *some* kind of hypothetical demand can ground a second-personal reason.



burdens are essentially suffered from a particular *temporal* perspective.<sup>20</sup> That is, the fact that the sufferer has a special perspective on the burdensome occurrence—a perspective characteristically had only from the first-person present standpoint—is essential to its being a burden at all.<sup>21</sup> Paradigmatic examples of such burdens are momentary pains or discomforts. When an individual receives an inoculation at the hospital, the needle's pinch counts as a burden precisely because of how things seem or how they are experienced from the (first-person present) perspective she has while the needle is inserted. If there weren't some way that the needle felt for a particular duration of time for the person getting the shot, then getting pricked by a needle would not count as a burden. Perspectives of this sort are continuously gained and lost as time passes. Because of this, we can make judgments about how well things go not just for an individual *simpliciter*, but also about how well they go for her from one temporal perspective or another. So we can say that things go fine from her perspective on her way to the doctor's office, badly from her perspective while the needle is inserted, a bit better from her perspective for the rest of the day as she nurses a sore arm, and then back to fine the next day.

With this, we are now in a position to see how one of our intra-personal cases parallels the inter-personal case in all respects relevant to the generation of second-personal reasons. The eighteen-year-old occupies a perspective from which smoking is pleasurable ( $P_1$ ), and so from which smoking is in her interest. The elderly person occupies a perspective from which the emphysema is painful ( $P_2$ ), and so from which it is burdensome. From  $P_1$ , the smoker could judge that she *would* demand that she quit smoking from  $P_2$ , and that she would be justified in so doing. Upon recognizing this, she could choose to regulate her action so as to conform to the

---

<sup>20</sup> Importantly, not all burdens are like this. If I hit my head and suffer amnesia, some think that my life goes worse because connections between the earlier and later perspectives are destroyed. This, despite there being no perspective at all from which that burden is suffered. (See Diane Jeske, "Persons, Compensation, and Utilitarianism," *The Philosophical Review* 100 (1993), pp. 541-575 (esp. pp. 558-560).)

<sup>21</sup> Arguably, the set of burdens essentially suffered from a temporal perspective is merely a subset of all the burdens suffered at a particular time. One might think, for instance, that an investor suffers significant burdens and, as a result, has a very bad day, even if she does not yet know that the market crashed and thus has no perspective on the event at all.

hypothetical demand, or she could choose to disregard it. Once again, we have the fact that there are distinct perspectives— $P_1$  and  $P_2$ —from which interests conflict. On the basis of the interests had from  $P_2$ , the individual could (hypothetically) make legitimate demands that she, from  $P_1$ , takes to be reason-giving. Thus, we have given a coherent account of what it would mean to take up the second-person standpoint with respect to oneself, and to generate second-personal reasons from that standpoint. And we did this without claiming that the wrongdoer and the victim are metaphysically distinct entities—we did not deny that X and Y are identical. Instead, we argued that a single person can herself be both the wrongdoer and the victim by virtue of having two different perspectives.

I have been focusing on wrongdoing involving welfare, but the argument applies equally well to the autonomy case. It is from a particular temporal perspective that an individual both endorses certain values and pursues them, and so it is from particular temporal perspectives that an individual exercises her autonomy. These perspectives perpetually come into and go out of existence, and an exercise of autonomy from one perspective might impinge upon the individual's ability to exercise her autonomy from another, later perspective. So then, to say that the nobleman owes it to himself not to constrain the fortune is simply to say that from his perspective as an older man a legitimate demand could be issued to not interfere with his ability to pursue the ends he will have as an older man.

Now let us revisit the Kant-Singer objection. On the account I sketched above, second-personal self-regarding reasons are generated *from a particular temporal perspective*. It stands to reason that if the individual were to successfully release herself from the duty, she would have to do so from the standpoint from which the (hypothetical) demand that generated the reason was issued. That is to say, since the (hypothetical) demand that generated the obligation not to smoke is issued from  $P_2$ , it is from  $P_2$  that she would have to release herself from the obligation. And, in the nobleman case, since the demand that generated the obligation not to constrain the fortune is issued from the perspective from which his autonomy is constrained, it is from *that*

perspective that the nobleman would have to release himself from the obligation. Both in the case of the smoker and in the case of the nobleman, the offending act occurs before the individual comes to have the perspective from which she could release herself from the duty. Securing the release is thus impossible. So self-regarding reasons can bind an individual in the way characteristic of moral duties after all. Kant and Singer's objection is thus answered.

#### 4. Duties to the Self

I have highlighted similarities between the inter-personal and intra-personal cases in order to establish the coherence of the notion of a duty to the self. However, there exist significant dissimilarities between the inter-personal and intra-personal cases, and one might suspect that some of these dissimilarities will provide justification for believing that an individual has duties to others, but not to herself. With respect to cases involving welfare, for instance, one might think that compensation occurs in the intra-personal case, whereas no compensation occurs in the inter-personal case.<sup>22</sup> So whereas the smoker is arguably compensated for her eventual suffering through the pleasure that smoking afforded her throughout the years, the community members who need the drug, or the person whose foot is trampled upon, receive no compensation when some other individual makes a profit or has her desires satisfied. Intuitively, this seems like a relevant difference. With respect to cases involving autonomy, the individual's choice in the intra-personal case, unlike the inter-personal case, results from an exercise of her own autonomy. So although the Russian nobleman constrains his future choice, he does so by making choices in accord with his own values. They are, for sure, values endorsed from a particular temporal perspective. But they are nevertheless *his* values. This seems relevantly different from an individual having her choice situation constrained by another's choice, made in accordance with that other individual's values—as happens when the Muslim is

---

<sup>22</sup> Rawls, *A Theory of Justice*, pp. 19-24; Nozick, *Anarchy, State, and Utopia*, pp. 32-33.

prevented by a Christian from building her mosque. In this section, I will argue that in spite of these differences, an individual nevertheless has duties to herself.

People do not merely care about, or have reason to concern themselves with, their lives going well “overall.” They also care about their lives going well from the present perspective. So it matters whether an individual is in pain *now*, or suffers humiliation *now*—even if she is fully compensated at some other time.<sup>23</sup> Indeed, it would get things exactly backwards to say of many burdens that an individual’s suffering them in the present is bad merely by virtue of their detracting from her life’s going well overall. For instance, if an individual experiences excruciating pain without incurring residual harms, the most natural thing to say is that her life goes slightly worse *because* things went badly from the momentary perspective. This is why we find our eighteen-year-old smoker intelligible, if misguided. By making the present moment pleasurable, she promotes something of genuine importance, even if she over-values it relative to other things worth promoting. Because it matters how well things go from the present perspective, there is genuine conflict between the value of things going well from the present perspective and the value of things going well from other perspectives (perspectives that will, at some point, be the individual’s present perspective). And similarly, there is genuine conflict between the value of things going well from the present perspective and the value of things going well for the individual overall, over the course of her whole life. For this reason, an individual will face a legitimate question about how to distribute goods over the course of her life. This, I will argue, opens the door for the individual to wrong herself, even if by imposing a burden upon

---

<sup>23</sup> Thomas Nagel attributes to Bernard Williams the observation that when someone says, “Thank God that’s over!”, she must be expressing concern for how well things go from the present perspective, and not expressing concern for the quality of her life on the whole. (See Thomas Nagel, *The Possibility of Altruism* (Princeton: Princeton University Press, 1970), p. 71 (fn. 1).)

herself she secures a future benefit sufficient to fully compensate her over the course of her life.<sup>24</sup>

On T. M. Scanlon's contractualist account, to act morally impermissibly is to act on a principle that could be reasonably rejected from some individual's "standpoint." Scanlon elaborates, writing that: "In order to decide whether a principle could be reasonably rejected, we need to consider it from a number of standpoints. From the point of view of those who will be its main beneficiaries, there may be strong generic reasons to insist on the principle and to reject anything that offers less. From the point of view of the agents who will be constrained by it, or of those who would be beneficiaries of an alternative principle, there may be reason to reject it in favor of something different or less demanding."<sup>25</sup> The idea here, appealed to in one form or another by everyone who emphasizes morality's second-personal character, is that objections made from the standpoint from which the burden is suffered can generate moral requirements and duties.

So now consider whether the smoker herself could reasonably reject the principle upon which she acts when she is eighteen, which appears to be: "Do whatever brings you the most pleasure now, come what may over the long run." To answer this question, the Scanlonian contractualist must consider the principle "from her standpoint." There is an ambiguity here, however. There might well be something we call "the standpoint of the person overall." But nevertheless, as we have noted, the individual's life will also contain multiple temporal perspectives—a succession of perspectives from which benefits can be incurred and burdens can be suffered. So there isn't merely a question about whether the individual can reasonably reject

---

<sup>24</sup> Many authors have recognized both that we care not just about how well things go for us overall, but also about how they go from particular perspectives, and that this can give rise to distributional concerns. For instance, Derek Parfit, after casting doubt upon the existence of a metaphysically deep self that persists through time, considers whether Rawlsian principles of distributive justice should be applied intra-personally rather than just inter-personally. (See his *Reasons and Persons*, pp. 329-339). Christine Korsgaard, who is not committed to Parfit's eliminativist view about the self, nevertheless raises similar questions about intra-personal distribution. (See her "The Myth of Egoism," in *The Constitution of Agency* (Oxford: Oxford University Press, 2008), pp. 69-99 (esp. p. 72).)

<sup>25</sup> Scanlon, *What We Owe to Each Other*, p. 213.

the principle, but also a question about which standpoints she might be able to do so from—from her standpoint overall, or, possibly, from one of the individual temporal standpoints along the way. The obvious candidate to consider first is the standpoint she occupies as she suffers from emphysema. It seems to me that from that perspective she has the same grounds for rejecting the principle that a person would have for rejecting a principle that allowed other individuals to foist significant burdens onto her: that the principle would license horrible suffering from her perspective. In fact, she might be able to reasonably reject a more reasonable-sounding principle, which says, “Do whatever will lead to the most pleasure, in aggregate, over the course of your life.” While she—the whole person—might be compensated for any burdens she would bear, she isn’t necessarily compensated from the perspective from which she suffers, and might reasonably reject such a principle on this basis.<sup>26</sup>

The Russian nobleman case can be handled similarly. Just as it matters whether an individual suffers burdens from a temporal perspective, it matters whether she is able to exercise her autonomy at particular moments. This is why, I take it, we feel a bit uncomfortable with the nobleman constraining his future self in the way he does in the example. This also explains why most liberals deny that a person may permissibly choose to sell herself into slavery—deny, that is, that an individual may choose now to strip herself of her autonomy in the future. On the basis of this concern for one’s ability to exercise autonomy, a principle that allows an individual to undermine her own ongoing autonomy could, depending upon the circumstance, be reasonably rejected from the perspective from which the lack of autonomy would be experienced. Even if the person herself decides, on the basis of her own values, to sell herself into slavery, it is nevertheless likely that she would be reasonable in rejecting a principle

---

<sup>26</sup> Importantly, though, this needn’t imply that it would be impermissible to impose *any* burdens on her future self. After all, she might be able to reasonably reject from the earlier perspective a principle so onerous that it prevents her from making any tradeoffs within her life whatsoever. My point here is simply that it is plausible to suggest that some burdens would be so great from a temporal perspective that it would be reasonable to reject the principle that licensed them, even if the principle promoted the individual’s overall good. If it would be reasonable to reject a principle that imposed such a burden, then a moral reason to refrain from imposing the burden has been generated.

that would allow her to do this—rejecting it, that is, from the perspective from which she would suffer from a lack of autonomy.

### 5. The Significance of Moral Reasons

I will conclude this chapter by explaining the significance of an individual's relating to herself second-personally. Many believe that there is a particular kind of practical necessity that comes with a consideration's being a "moral" one—they talk as if moral considerations particularly weighty, or stringent. Non-moral reasons are, on the other hand, thought to be less stringent. For instance, Wallace says that despite having decisive reason to attend the screening of *Rocco and his Brothers*, our potential moviegoer intuitively has, "a little leeway to ignore or discount the considerations that speak in favor of [attending]." <sup>27</sup> An individual has, in this instance, the prerogative to be irrational—the prerogative to not do as she ought. The loan case is different, however. There Wallace says that we, "lack discretion to ignore or discount the reasons [that favor repaying the loan] in a way that makes it natural to think of them as requirements." <sup>28</sup> Wallace thinks that what gives moral reasons their stringent, non-optional character is the fact that there is an agent in a position to legitimately demand the repayment. Basically, because a person owes *to someone* the performance of an act, she lacks the authority to decide not to act. Whereas the individual might whimsically do what she knows is irrational on some occasions, the prerogative to do so is lacking when another person has the authority to demand that one act. And to relate this to our political topic, this is why the state is arguably so severely constrained when it makes decisions. As it makes policy, it risks harming its members, drawing them into its pursuits, and undermining their own projects—and many think that it lacks the authority, or the discretion, to do these things.

---

<sup>27</sup> Wallace, "The Deontic Structure of Morality," p. 8.

<sup>28</sup> Ibid., p. 12.

If this is right, though, then the individual and the commonwealth have something important in common. When they act, they both are constrained in a similar way—by reasons that they lack the authority to disregard. So if the individual and the commonwealth differ greatly in terms of the kinds of things they may permissibly do, the difference does not lie in the individual's lacking obligations to herself. If, as most believe, an individual really does, in the end, have extensive moral permission to pursue her own good and her own ends, an explanation for this is required. We must explain why it is that she may permissibly do these things, given the obligations she has to herself. And once the explanation is given, we can then use the explanation as a model for constructing an explanation of why the commonwealth may permissibly impose burdens on citizens in the pursuit of the overall good, as well as pursue its collective ends.



## Chapter 4: The Common Good

### 1. Introduction

In Chapter 2, I argued that a commonwealth is an agent that is under normative pressure to find good or worthwhile things to do, and then do them. However, significant questions remain about what it is permissible for the commonwealth to do—about the ways in which it will be morally constrained as it attempts to carry out good or worthwhile action. After all, we cannot, of course, simply assume that whatever it does in the process of carrying out its worthwhile endeavors is permissible. In this chapter, I will consider whether a commonwealth may permissibly impose burdens on some of its members as it realizes its overall good.

An *individual* agent attempting to live a life that is worthwhile will often make tradeoffs, imposing burdens upon herself for the sake of promoting her overall good. She might compete as an Olympian, knowing the strain on her body will require painful surgery. She might attend an expensive art school, knowing the debt she accrues will force her to work long hours later on. She might invest significant efforts into a rewarding endeavor early in life (such as an academic project or a small business), knowing there is a significant risk of failure for which she will bear the cost later on. While living a worthwhile life does not *only* involve making these sorts of tradeoffs, making such tradeoffs certainly constitutes a significant portion of it. But as I argued in the previous chapter, it is possible for a person to wrong herself by burdening herself, and so we cannot simply assume that it is permissible for her to make tradeoffs. We need an argument that shows that the individual does not wrong herself by imposing burdens upon herself. So I will begin this chapter by arguing that an individual is often permitted to make tradeoffs because she is compensated through contributions to her overall good—indeed, compensated *at the very moment that she bears the burden*. Then, in the rest of the chapter, I will oppose the dominant liberal view by arguing that the commonwealth may similarly pursue the common

good by imposing burdens on individuals, because the realization of the common good compensates burdened members for the burdens they bear.

I will begin in section 2 by discussing compensation and its role in moral and political justification. In section 3, I discuss how compensation occurs within an individual's life, with an eye to gaining insight that can then be applied to the commonwealth. In sections 4-6, I argue that group action, carried out by an entire commonwealth, constitutes a common good in which all individual members can share. This, I argue, enables compensation to occur, even when an individual is burdened in the pursuit of the common good. In section 7, I answer some worries about the common good's association with totalitarianism.

## 2. Compensation and Justification

While a burdened individual may sometimes be wronged despite being compensated, in some cases the imposition of otherwise wrongful burdens is made justifiable by virtue of the fact that compensation occurs.<sup>1</sup> For instance, implementation of public policy inevitably burdens someone, whether through a requirement to pay a tax that diminishes an individual's wealth, a regulation that makes doing business more expensive, or a traffic law that makes a commute more time consuming. But it is plausible to think that each individual has a moral claim, in the policy-making process, to have her interests respected. When a particular policy burdens an individual, one way to render the policy justifiable to her might be to eventually bestow a benefit upon her, so that she is made better off overall. The state might, for example, allow the taxpayer to use the infrastructure built with her tax dollars, allow her to take a large deduction on her taxes to compensate for the regulatory burden, or point out that the traffic law will eventually result in her paying a significantly lower insurance premium. In each example, the fact that the

---

<sup>1</sup> Seana Shiffrin gives an example of a person who drops gold bars out of an airplane in order to benefit the people below, and breaks one of their arms. While the person is undoubtedly better off for having the gold bar, it still seems like the person was wronged by having her body subject to an injury to which she did not consent. All this shows, however, is that the fact that an individual is compensated is not *sufficient* to justify the imposition of a burden. (See her "Wrongful Life, Procreative Responsibility, and the Significance of Harm," *Legal Theory*, 5 (1999), pp. 117-148, (esp. p. 127).)

individual is eventually compensated seems relevant to the question of whether she was wronged by having to bear the burden.

While these claims about compensation within an individual's life and the role that compensation can play in moral justification are widely accepted, analogous claims about entire societies or communities are almost universally rejected. Robert Nozick explains:

Individually, we each sometimes choose to undergo some pain or sacrifice for a greater benefit or to avoid a greater harm: we go to the dentist to avoid worse suffering later; we do some unpleasant work for its results; some persons diet to improve their health or good looks; some save money to support themselves when they are older. In each case, some cost is borne for the sake of the greater overall good. Why not, *similarly*, hold that some persons have to bear some costs that benefit other persons more, for the sake of the overall social good? But there is no *social entity* that undergoes some sacrifice for its own good. There are only individual people, different individual people, with their own individual lives.<sup>2</sup>

Nozick's claim that there is "no social entity that undergoes some sacrifice for its own good" is a bit obscure, but the basic point seems to be that while individual persons exist in some metaphysically deep sense, "social entities" do not. Social entities consist merely of individuals and the relations between them, and nothing about this suggests that the social entity itself could suffer a harm or be compensated. Instead, what happens when a burden is imposed in the pursuit of society's overall good is that some individuals are harmed and are not themselves compensated.<sup>3</sup> We therefore cannot appeal to compensation in such instances in order to justify imposing a burden on an individual. If Nozick is right, this is a significant difference between

---

<sup>2</sup> Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 32-33.

<sup>3</sup> Thomas Nagel makes essentially the same point, claiming that it is morally significant that while a single person may be compensated later for burdens borne in the present, inter-personal compensation is not possible. (See his *The Possibility of Altruism* (Princeton: Princeton University Press, 1970.), pp. 138-142.)

the individual and the commonwealth. Whereas the individual is permitted to make tradeoffs within her life in order to pursue her own overall good, the commonwealth is not.

While Nozick is attempting to highlight a way in which society and individual persons differ, Derek Parfit gives a theory of the person on which the individual resembles a Nozickian society. Against thinkers such as Geach, Chisholm, and Swinburne, Parfit argues that there is no deep metaphysical fact about whether an individual is the same person from one moment to the next—that there is no “self” over and above psychological states at particular moments and the connections between them.<sup>4</sup> Parfit calls this his “Reductionism” about personal identity. So, just as Nozick denies the existence of a social entity over and above the individual members and the relations between them, Parfit denies the existence of a self over and above psychological states and the connections between them. And this, Parfit thinks, raises doubts about whether compensation can occur within a life:

[Because there is no deep fact about whether one exists from one moment to the next, one] can defensibly claim that a benefit at one time cannot provide compensation for a burden at another time, even when both come within the same life. There can only be simultaneous compensation, as when exposing my face to a freezing wind is fully compensated by the sight of the sublime view from the mountain I have climbed.<sup>5</sup>

Parfit thus raises a doubt about the permissibility of an individual's burdening her future self that mirror's Nozick's doubt about the permissibility of the commonwealth's burdening an individual citizen.

---

<sup>4</sup> Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), pp. 261-266. Hume had a similarly reductive view of the self. He compares the mind to, "a system of different perceptions or different existences, which are linked together by the relation of cause and effect." For this reason, Hume says, "I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts." (*A Treatise of Human Nature 2<sup>nd</sup> Edition*, ed. L.A. Selby-Bigge (USA: Oxford University Press, 1978), p. 261 (Book I.IV.4).)

<sup>5</sup> Parfit, *Reasons and Persons*, p. 343.

So in Parfit and Nozick, we have examples of the prominent belief that important moral issues hinge on whether there is a deep metaphysical fact about the existence of a self, or a deep metaphysical fact about the existence of a social entity. Against this, I will first argue that compensation within a life is possible, and that this often makes it permissible for an individual to burden herself in pursuit of her overall good, *regardless of whether there is such thing as a metaphysically deep self*. I will then use this account as a model for the argument I will make later in the chapter, which addresses Nozick's worry. I will argue that an individual can be compensated through the advancement of the common good, and that this often makes it permissible for the commonwealth to burden the individual in pursuit of the common good, *regardless of whether there is such thing as a metaphysically deep social entity*.

### 3. Compensation and the Unity of Action

In the previous chapter, I argued that duties to the self can be generated from a particular temporal perspective. The fact that they can be is explained, in part, by the fact that some burdens are suffered from a particular temporal perspective. Momentary pain was our paradigmatic example. It is plausible, we concluded, to think that the individual could reasonably reject, from the perspective from which she would suffer, the principle of action that would permit her to impose severe burdens upon herself. Thus, she might sometimes have a duty to avoid burdening herself, even if the imposition of the burden were necessary in order to make her life go better overall. This account, however, casts doubt upon whether compensation can play any role in justifying tradeoffs within a life. Even if a person persists through time in some "metaphysically deep" sense, and for that reason is able to both bear the burdens and incur the benefits, the issue is not simply whether the whole person is ultimately compensated. If the benefit is not enjoyed from the perspective from which the burden is suffered, then the benefit cannot help to justify the burden by making the individual better off from the perspective from which she issued the demand that generated the duty. That is, assuming the

demand that generated the duty is a demand to give a certain regard to the interests had from a particular temporal perspective, bestowing a benefit that will be enjoyed from an entirely different temporal perspective will not address that specific demand.

Earlier I quoted Parfit, who claimed that the only possible compensation is *simultaneous* compensation. His example was someone who simultaneously endures cold wind in her face in order to enjoy a beautiful view on top of a mountain. Because there is a *single perspective* from which both the benefits and burdens are incurred, compensation is possible, Parfit thinks. In this section, I will demonstrate that individuals who accept a momentary burden for the sake of advancing their overall good are sometimes importantly similar to Parfit's mountain climber. Specifically, I will argue that an individual can sometimes be compensated *at this very moment*, or *from her momentary standpoint*, by goods that seem to be bestowed outside of the present.

My argument will rest on the claim that part of an individual's good consists in engaging in worthwhile activity, so I want to say a bit about this first. While pleasurable sensations or experiences are obvious constituents of an individual's good, passive experiences are only part of what constitutes a human life. A life is also constituted by one's action, and so what an individual *does* is a sensible dimension along which to evaluate how well her life goes. This is a point made by Aristotle, most clearly in Book X of the *Nicomachean Ethics*. There Aristotle argues that pleasure alone cannot be the sole constituent of the human good, as carrying out bad or vicious activity, while all the while enjoying it, makes a person's life go worse, not better.<sup>6</sup> In fact, if a person spends her days tormenting defenseless animals, or cheating others out of their savings, the fact that she enjoys these things so greatly is even more of a reason to think her life goes badly. Now, our intuitions might pull us in opposite directions here, as it can easily seem

---

<sup>6</sup> Aristotle, *Nicomachean Ethics*, 1173<sup>a</sup> 14 - 1174<sup>b</sup> 10.

that the individual enjoys *some* benefit when she takes pleasure in her vicious activity.<sup>7</sup> But we might see the pull of Aristotle's point by thinking about what we would say about a loved one—an individual whose welfare we care about—who engages in vicious acts.<sup>8</sup> Imagine, for instance, that a loved one begins behaving hatefully toward members of another race. One might encourage her to quit acting this way for a variety of reasons, including, first and foremost, concern over wrong done to her victims. But one might also discourage the loved one from acting this way out of concern for *her* well-being—out of concern that her action itself makes her life go worse. Even if she thoroughly enjoyed degrading others in the way she does, and would likely not enjoy anything else nearly as much, it still seems sensible to encourage her to stop doing so, not only out of concern for the victims, but out of concern *for her*. This seems plausible, I believe, because we accept that good action constitutes part of an individual's good.

The fact that worthwhile action can constitute part of an individual's good explains how goods that may seem to accrue outside of the present can infect the present. Or so I will argue. To begin, consider that much of the activity that an individual is engaged in at any particular moment is constitutive of some larger action. We often say of an individual who acts at a particular moment: "By doing A, she is  $\phi$ -ing" (where  $\phi$ -ing is some temporally extended activity that unfolds over time). Michael Thompson explains:

[S]ome of the temporally extended intentional actions that interest us are . . . intuitively resolvable into a heterogeneous collection of sub-actions that are themselves clearly intentional—organs, as it were, of the whole. Such is the relation of egg-breaking and egg-mixing to omelet-making, of

---

<sup>7</sup> John McDowell, though, has suggested that sacrifice for the sake of virtue is really not sacrifice. (See his, "The Role of *Eudaimonia* in Aristotle's Ethics," in *Mind, Value, and Reality* (Cambridge, Mass.: Harvard University Press, 1998), pp. 3-22.)

<sup>8</sup> See Philippa Foot, *Natural Goodness* (Oxford: Oxford University Press, 2001), pp. 93-98.

brick-laying and door-framing to house-building, and of writing the letters “a” and “c” to writing the word “action.”<sup>9</sup>

So when an individual cracks an egg on an occasion in which she is omelet-making, her action is not distinct from the omelet-making itself. "Making an omelet" is every bit as true a description of what she is doing as "cracking an egg." However, when the individual adds some vegetables to the mix a few moments later, she is also making the omelet—the very same omelet. Thus, she is carrying out the very same activity or action.

It appears, then, that we can say: "In doing A, she is  $\phi$ -ing, and in doing B a few moments later, she is also  $\phi$ -ing." So an act performed at a particular moment (A) forms a kind of unity with goings-on outside of the present (B). A and B are unified by their both being phases of  $\phi$ -ing. The fact that  $\phi$ -ing spans a period of time enables it to unite the future with the present—to unite A and B. Let us now stipulate that  $\phi$ -ing is worthwhile, but that what makes it worthwhile is more apparent when considering some of its constitutive phases rather than others—for instance, when we think of what is worthwhile about  $\phi$ -ing, we are likely to appeal to B but not A. Or, alternatively, suppose that we appeal to the way that the phases of  $\phi$ -ing interact in order to explain why  $\phi$ -ing is worthwhile. We might then appeal to the fact that in doing A, one is  $\phi$ -ing, in order to explain how doing A is beneficial to her—in order to explain how doing A advances her interests.

To illustrate, imagine that an individual cracks an egg. But in so doing, she is making an omelet. And in doing *that*, she is carrying on with her career as a master chef. Which, since she is a genuine talent, is an exercise of her skill—a realization of her potential. So our chef, as she cracks the egg, is doing something much more profound than what a child does when she frivolously cracks an egg for the sheer joy of seeing it break. The chef is carrying on with a particular kind of career and is living a particular kind of life—one in which she flourishes by

---

<sup>9</sup> Michael Thompson, "Naïve Action Theory," in *Life and Action: Elementary Structures of Practice and Practical Thought* (Cambridge, Mass.: Harvard University Press, 2008), pp. 85-119 (esp. p. 106).



making use of her talents. I believe we can say, with plausibility, that action that constitutes a realization of one's potential is action worth doing. If this is right, then the chef who gets out the eggs and cracks them is doing quite well for herself at that very moment. After all, we have already established, in the previous chapter, that we can evaluate how well things go for an individual from a temporal perspective. So if we allow that part of the human good is constituted by worthwhile action, then it stands to reason that the person does well from her perspective when she carries out, from that perspective, a worthwhile action. But, importantly, in the case we are considering presently, this is not because there is anything worthwhile about the egg-cracking per se. It is worthwhile because of the temporally extended life-activity of which egg-cracking is an organ. The individual's interests are advanced not simply because she does A, but because in doing A, she is  $\phi$ -ing.

The way that ongoing action enables the future to infect the present makes compensation possible. Even if doing A, considered in isolation, is burdensome, the individual might be compensated from the perspective from which she does A by the fact that she is  $\phi$ -ing. To illustrate, if our master chef is making risotto, she will be doing a lot of continuous stirring over a hot stove, which might well be burdensome. But, as the chef stirs, she is not doing something distinct from risotto-making. As she stirs, she is *making the risotto*. By describing what she is doing in this way, the parallels with Parfit's mountain-climber scenario—where we were confident that compensation occurred—begin to emerge. From the chef's momentary perspective, risotto-making is in her interest, and stirring continuously is a burden. At this particular moment, however, the two are *the very same action*, done *from the same momentary perspective*. The chef thus cannot avoid the tedious stirring and still engage in risotto-making, any more than Parfit's mountain climber can avoid the cold wind while enjoying the view. As the chef endures the stirring, she is simultaneously compensated by the fact that she is flourishing by practicing her craft.

The basic account here can be applied to a wide range of cases. Imagine that a parent sits in traffic on her way to her daughter's elementary school play. This is a burden. But this—the sitting in traffic in order to get to the play—is all part of being a good parent. So while the individual bears a burden from her momentary perspective as she sits in traffic, she also benefits from that same perspective by carrying on with her parenting (and doing it well). Thus, she is compensated. Or, imagine that a student sits in a dark library, reading a dull book. This is a burden. But in so doing, she carries out her larger project of becoming a lawyer and defending the poor and dispossessed. Thus, the good of participating in this project compensates her for the burden at the moment she suffers it. Or, imagine that an aspiring political leader solicits donations. This is unpleasant and burdensome. But this is all part of her running for office and giving back to her community. By acting as a good citizen, which is worthwhile, she is compensated for having to do this unpleasant task from the perspective from which she carries it out. These examples have all involved burdensome acts that partially constitute worthwhile activities. But a similar point can be made about burdensome acts that simply *facilitate* worthwhile activity. For instance, going for a flu shot might be a practical demand of any number of worthwhile activities. Being a parent places a practical demand on one to (among many other things) look after the health of one's children, to transport them to school, and to attend their plays and sporting events. One's ability to perform these tasks is put in jeopardy when one has the flu. So the following bit of reasoning seems sensible: "I am carrying on as a parent. This requires I stay healthy. So I will get a flu shot." It is not natural to say that at the moment that the parent gets the shot, what she is doing is *constitutive* of parenting (in the way that what the stirring chef is doing is constitutive of risotto-making). But at the moment the parent gets the shot, it isn't as if she is no longer carrying on with her parenting. Indeed, by getting the shot, she is acting as a conscientious parent. So here too we might judge that carrying on as a conscientious parent compensates the individual for the needle prick at the very moment that she gets the shot.

It is thus possible for an individual to be compensated for burdens that she bears, from the perspective from which she bears them, through things that happen outside the present. This is possible even without positing the existence of a metaphysically deep self. It is for this reason that we can sometimes appeal to the fact of compensation when explaining why it is permissible to impose a burden that will be suffered from a particular perspective. I do not want to overstate things here, as the fact that an individual is compensated is probably neither necessary nor sufficient to justify imposition of a burden. But it can help explain, in a broad range of cases, why imposing burdens on oneself in order to do something worthwhile is morally permissible.

The point, then, is that by advancing one's good by engaging in worthwhile action, one is compensated for momentary burdens as they occur, making it plausible to say that an individual can realize her own good without worrying that the burdens she imposes upon herself in so doing constitute a moral wrong.

#### 4. The Common Good and Compensation

The argument above will serve as a model for my argument that compensation between distinct individuals within a commonwealth is possible. In the remainder of the chapter, I will argue that by engaging in good or worthwhile collective action, the commonwealth realizes the common good, and in so doing compensates citizens who are burdened by the actions necessary to realize it. By a "common good," I mean a state of affairs, process, or event that *constitutes* the good of each individual under consideration, as opposed to a state of affairs, process, or event that merely promotes some *other* interests that the individuals have. This distinction is a bit delicate, so I will pause to spell it out. Consider those things that we are inclined to say are simply "good for me," or "Me-goods." Here I have in mind my own material wealth, my health, my lack of discomfort or pain, freedom or means with which to pursue my ends, and so on. The goodness of such things is intimately connected to the contribution they make to my life's going better. A

“coincident good” is something that simultaneously promotes items that occupy a position on both my “Me-good” list, and on every other individual’s “Me-good” list. For instance, upon discovering a valuable reserve of some natural resource, the state might mine it and distribute the resulting wealth to each of its citizens. The policy is good for me because it increases *my* wealth, it is good for you because it increases *your* wealth, and it is good for Bob because it increases *Bob’s* wealth. This is different from a policy that brings about a particular state of affairs or event that *constitutes* the good of each person under consideration. This kind of policy would promote what I call a “common good.” So for instance, if Bob’s being healthy counted as making him better off, but also, in itself, counted as making me (along with everyone else) better off, then Bob’s health would be a common good. To put it loosely, to suggest that there exists a common good is to suggest that people’s interests can be intertwined rather than merely aligned.

I draw this distinction between a common good on the one hand and coincident goods on the other not because I believe that only common goods should be promoted. If, at little expense, the state could clean the Earth’s atmosphere and in so doing promote the health of each person—which would be to promote a number of coincident goods—then there is undeniably a strong case for doing so. Rather, I draw this distinction because I want it to be clear that my thesis is *not* that the commonwealth’s engaging in worthwhile group action will promote coincident goods—my thesis is not that engaging in worthwhile activity will make each person more wealthy, or more healthy, or more able to pursue her own ends. Rather, my view is that the commonwealth’s engaging in worthwhile group action *in itself* constitutes a good of each—it *in itself* makes a contribution to each individual’s life going better.

Now, recall the problem with which we began, which was that in enacting public policy, a commonwealth inevitably burdens some individuals in the process. Those individuals are owed a justification for having to bear the burden. If Nozick is right, we cannot appeal to the concept of compensation in order to give that justification. However, if we connect the concept of worthwhile group action to the concept of a common good—if we show that by engaging in the

former, we realize the latter—then we show that by engaging in worthwhile action, the commonwealth would promote the good of each. If some are burdened in the pursuit of that overall good, they will also be compensated by the achievement of that good.<sup>10</sup> So this connection between the concept of worthwhile group action and the concept of a common good is what I hope to establish.

## 5. Worthwhile Group Action as a Good

Philosophers have given many accounts of how people's interests can be intertwined or shared. The problem for our purposes here is that most accounts seem most applicable to the context of familial relations, friendships, and small communities, and not to citizens living together in large, modern societies.<sup>11</sup> Our question, then, is how citizens' goods or interests could be intertwined within such societies. As I said above, my hypothesis is that successfully performing good or worthwhile group actions together as a society constitutes an interest of each member, and thus is a common good. I will argue for this conclusion first by demonstrating that engaging in worthwhile group action generally can constitute a good for an individual, and then arguing that actions undertaken as a member of a large modern society are among those group actions that are likely to advance an individual's overall good.

To begin, recall my claim from section 3 that good or worthwhile action partially constitutes an individual's good. For reasons similar to the ones I gave there, I think that group

---

<sup>10</sup> Whether they are compensated *fully*, however, will depend on how great the burden is, and how great a good the particular common good is.

<sup>11</sup> For instance, T.M. Scanlon suggests that there is no need to identify sharp boundaries dividing one person's interests from another's. Need we, for instance, *explain* how a parent's interest is intertwined with his child's? Or might we instead allow that the interests of one person can simply seep out into others'? After all, if my child lives a happy and successful life, it seems perfectly intelligible to say that this *just does* contribute to my life's going well also. (See his *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998), pp. 126-129.) The reason why Scanlon's view fails to help us with our present question is that it isn't clear at all what the state could even do with this observation. Certainly the boundary between people's interests isn't so blurry that a benefit to one person could constitute an interest of all or even a significant portion of a population—at least if we are imagining a society constituted by a significant number of people. Whatever goods might be common due to a vague boundary between people's interests are not common enough to serve as a common good for an entire society.

action can constitute an individual's good. To illustrate the intuitive thought, imagine an expert mountain climber. It is plausible to judge that it would be in such a person's interest to climb Mount Kilimanjaro. The action is neither evil nor vicious. It requires exercise of the climber's talents. It allows her to succeed in an endeavor to which she has devoted herself. So it is sensible to say that her life goes better—that she advances her interests—if she climbs the mountain successfully. Now, satisfying this interest involves more than realizing a state of affairs in which she is at the summit. Were she dragged to the top while asleep, or transported there in a helicopter by kidnappers, the interest would not be satisfied. This is, of course, because her interest is satisfied only by performing an *action*—by moving herself for *reasons*. As I pointed out in Chapter 1, performing an action often involves doing one thing in order to do something else. So the individual engages in practical thinking of the form: "I am  $\phi$ -ing. Therefore, I will A." In the present example, then, the climber must engage in bits of thinking such as: "I am climbing this mountain. Therefore, I will take this next step," or: "I am climbing this mountain. Therefore, I will check my map."<sup>12</sup> It is only then that the interest that she has in climbing the mountain is fulfilled. I also argued, again in Chapter 1, that humans can engage in group action by carrying out acts that constitute a larger action, which other individuals also help to carry out. Individuals engage in practical thinking that has the form: "We are  $\phi$ -ing. Therefore, I will A." So, for instance, an individual might engage in bits of thinking such as: "We are climbing this mountain. Therefore, I will take this step," or: "We are climbing this mountain. Therefore, I will help my partner." The Aristotelian case for claiming that an individual's good is partially constituted by her action rested on the truism that a human life has an active aspect, in addition to a passive one. But an individual's capacity for activity is not exercised only in those instances in which she would say, "I am doing A, because I am  $\phi$ -ing." She is also exercising it in those

---

<sup>12</sup> I am imagining this as a rational reconstruction of an individual's thoughts, rather than that the agent explicitly says such things to himself.

instances in which she would say, "I am doing A, because We are  $\phi$ -ing." So when evaluating how well the active aspects of her life go, group action in which she partakes seems relevant.

Indeed, explaining why a particular action is in a person's interest at all might well require making reference to the fact that it is a group action. As I explained in section 3, in many cases, the good of doing A is explained by the fact that in doing A, one is  $\phi$ -ing. That is to say, the good of doing A is parasitic of the good of  $\phi$ -ing. To understand why it was in the interest of the chef to crack the egg, one needed to understand why it was in her interest to be making an omelet. The same point holds for group action. When engaging in group action, an individual takes the fact that *We* are doing something as her reason for acting. For instance, she reasons: "We are climbing the mountain. Therefore, I am lifting my foot half way, and now the rest of the way, and now I am moving it forward and down half way, then the rest of the way, and I am checking to see if my partner's foot is moving, and now. . ." We certainly might imagine someone who, without deploying any group action concepts, thinks: "I will lift my foot half way, and now the rest of the way, and I will now move it forward and down half way, then the rest of the way, and I will check to see if my partner's foot is moving, and now. . ." But it would be entirely unclear why performing *this* action would advance the individual's interest. This is because if it is in an individual's interest to carry out such an action, it will usually be because in doing so, she will be performing an action whose unity we can appeal to in order to explain why the act is worth doing. The good of doing A is often explained by the  $\phi$ -ing that it constitutes, and sometimes the  $\phi$ -ing is an action carried out by a group.

The challenge now is to apply the conclusions arrived at above to groups the size of an entire modern society. One might doubt, after all, that cases like "Our climbing the mountain" are perfect models for understanding how society-wide action can advance an individual's good. While people set out to climb mountains with partners, they rarely set out *simply* to contribute to society's projects. Something more complicated thus needs to be said about society-wide action.

To begin, imagine a person who can live under one of two societal arrangements, each of which affords her the same career opportunities and earning potential. In the first, which I will call "Museum," the individual's activity—her labor—is harnessed through a system of taxation so as to create a first-rate museum system that enriches the general public. In the second, which I will call "Star Wars," the individual's activity is harnessed through a system of taxation so as to advance an ill-advised quixotic missile defense system, which burns through money without showing tangible results. Whichever arrangement she lives under, the individual, in carrying on with her career and paying her taxes, contributes to the project. So when policymakers adopt a plan, their practical thinking will take the form: "We will  $\phi$ , and will carry this out by harnessing the efforts of individuals, working their jobs." For instance, they might say: "We will pursue Star Wars, and we will carry this out by Bob's running a small business, and Peter's working at the law practice, and Bill's working at the auto plant." About each individual, then, we can make a judgment of the form, "by her doing A, we are  $\phi$ -ing." So we can judge, "by running her small business, the society maintains the museum system," or, "by her working as a lawyer, the society pursues the Star Wars program."

The individual is thus, through the decisions of policymakers, drawn into a project. This is true even if we never judge of the individual that she, as an individual agent, does her job *in order to* make a contribution to the museums or to Star Wars. We often say, "in doing A, the individual  $\phi$ 's" without saying, "The individual does A in order to  $\phi$ "—indeed, we do so quite often. Consider, for instance, that loving acts are often constitutive of maintaining and carrying on with a loving relationship. Maintaining a relationship consists partly of such acts, and maintenance of the relationship is attributable to the person by virtue of the fact that she engages in such acts. But momentary loving acts are usually not done *in order to* maintain the relationship. They are done for their own sake, or for pleasure, or to make another person happy. The same can be said of acts that constitute destroying a relationship. An individual might destroy her relationship with her spouse by carrying on an affair without carrying on the



affair *in order to* destroy her relationship. Similarly, a contribution to the Star Wars project might be attributable to the taxed worker who does her job simply because she needs to make a living, or because she finds her career fulfilling, but who does not labor *in order to* support Star Wars.

The fact that an individual's action can be redescribed in terms of another action that it constitutes—the fact that in doing A, the individual  $\phi$ 's—is often relevant when assessing how well the individual's life goes. Oedipus knowingly kills a stranger at the crossroads, on his way to Thebes. It is relevant to our assessment of how his life goes that in so doing, he fulfills the Oracle's prophecy and kills his father. Oedipus marries Jocasta. It is relevant to our assessment of how his life goes that in so doing, he fulfills the Oracle's prophecy and sleeps with his mother. The tragedy of the Oedipus story consists in the fact that his actions, when considered under those descriptions, are calamitous.

What about, then, action carried about by an entire society? To give an illustration involving participation in society-wide action, consider now the plot of Kurt Vonnegut's novel *Mother Night*.<sup>13</sup> In the novel, an American playwright named Howard Campbell finds himself living in Germany just prior to the outbreak of World War II. Campbell is approached by the U.S. State Department and is asked to serve as a spy during the impending war—specifically he is asked to relay information through coded messages while acting as a Nazi propagandist, broadcasting on German radio. Campbell accepts the mission, not so much out of a desire to aid the Allied Forces, but mostly because he finds this to be an innocuous way of remaining in Germany where he has been living. So he spends years carrying out this task, spewing vile Nazi propaganda, and in so doing, supports the Allied effort. When all is said and done, there are several ways to describe what he did. On one description, he was just biding his time until the war ended. On another, he helped the Allies defeat the Nazis. On yet another, he spent every day giving legitimacy to some of the most heinous ideas imaginable. The book invites us to ask what

---

<sup>13</sup> Kurt Vonnegut, *Mother Night* (New York: Dell Publishing, 1961).

kind of life Campbell lived, and all the descriptions seem relevant. The fact that he bides his time seems relatively innocent. The fact that the way he did so can be redescribed as "propagating Nazi ideology" indicates that there is much room for improvement. Campbell, it would seem, would have lived better had this not been a true description of how he spent his time. But we also seem to miss something relevant if we ignore the fact that what he did can also be redescribed in terms of the cooperative Allied effort. We would feel differently about Campbell and the type of life that he lived had he become a Nazi propagandist in order to remain in Germany without simultaneously participating in the Allied effort.

The question about whether Campbell lived a good life has no obvious answer. What is obvious is that when judging how well he lived, how his activity can be redescribed is often not only relevant, but *crucial*. So I think we can conclude that an agent's life goes better if her activity can be redescribed in terms that show it to be constitutive of worthwhile group action, rather than constitutive of some less valuable group action, or some feeble group action, or some bad group action, or, perhaps, no group action at all.

Return now to "Museum" and "Star Wars." Since I have stipulated that the missile defense system is a waste, the individual who engages in producing it, even if her contribution to the effort takes the circuitous route of paying taxes, will necessarily be engaged in a wasteful activity. Though there might be many descriptions under which what the individual does is valuable—her work might be engaging in its own right, for instance—it is certainly true that there is room to improve the activity, at least along one dimension. Given this, so long as we admit that the museums have *some* value because they enrich the lives of the people who visit them, we should be prepared to admit that one's interests are advanced if, in doing one's job, one works to realize the museum system rather than the worthless missile defense system.

Now, I began with "Museum" because it is relatively easy to see how building and maintaining a museum would be good or worthwhile. An action that produces something of value is, itself, usually good or worthwhile by virtue of this fact. But we can move on to ask more

generally which standards we ought to use to evaluate any particular group action's worth. Perhaps group action that exhibits certain human excellences contributes to the individual's life's going well. Or perhaps group action that is virtuous does. Or perhaps group action that is *just* does. We can debate about the appropriate standards to apply, but society as a whole could manage to engage in activity that either met or violated standards along these lines—suggesting that society's activity could, in itself, constitute a contribution to or a detraction from its members' good, even in instances where the activity isn't one that produces something.

## 6. The “Intertwining” of Interests

In this section, I will explain how the account given above amounts to saying that individuals' interests can be “intertwined.” If by my doing A, and by your doing B, We  $\phi$ , then my doing A and your doing B form a unity. “We are  $\phi$ -ing” is a true description of what both you and I do. If in my doing A, and in your doing B, We are  $\phi$ -ing, and Our  $\phi$ -ing constitutes an interest of each of ours, then your doing B in itself contributes to my good. Indeed, assuming that an interest in  $\phi$ -ing is an interest in  $\phi$ -ing *successfully*, your doing B successfully will contribute to my good, while your doing B *unsuccessfully* will undermine my good. This is what I mean when I talk of interests being intertwined through group action. So returning to “Museum,” if the creation and maintenance of something valuable is in each of our interests, then things that undermine its creation and maintenance make each of our lives go a bit worse. If the whole museum system ultimately goes bankrupt and fails, then we are all made worse off by virtue of the fact that our worthwhile collective activity failed. If some citizens undermine the activity or contribute to its going less well than it might have—either by not paying their full share of taxes, or by littering or otherwise treating the museums disrespectfully when visiting—then each is a bit worse off for having the collective effort undermined. So I think it is relatively clear how through worthwhile group projects of this kind, the interests of all the citizens' are intertwined.

But now it might seem like I have pulled a bit of a bait and switch. Earlier, while defining the concept of a common good, I discussed the possibility that one person's health or wealth could constitute the good of another. Now, however, it seems that I have quit talking about one person's health or wealth constituting the good of another, and have instead started talking about how building museums could constitute everyone's good—an interesting conclusion, perhaps, but less ambitious and more circumscribed than the initial hoped-for result.

While it is true that my view has not explained how the state of affairs in which one person is healthy or wealthy can constitute the good of another, I think there is a way to use my view to support conclusions that are equally significant. To use the example of health, an entire people might implement and sustain a particular health care system. Implementing and sustaining the system is a group activity to which the entire society contributes, and thus is a candidate for constituting the good of each in the sense I have described. What seems to be left out of the story, though, is anything about the health of the individuals who benefit from the health care system. But it cannot actually be true that people's health has been left out of the story, at least if we are assuming that sustaining the system actually succeeds in constituting the good of each. After all, if the activity does constitute the good of those carrying it out, it must be because the activity is good or worthwhile. And if this particular activity—our sustaining a health care system that provides care for all—is indeed worthwhile, its value must have *something* to do with the health that the activity promotes or sustains. So while the health of one person does not, in the end, constitute the good of another, the health of one person bestows value on an activity that, because of this value, constitutes the good of each participant.

This general strategy can easily be redeployed across a wide range of cases: A system of public universities provides something educational for anyone who wants to take advantage of it. This fact, in turn, explains why maintaining a public university system is in everyone's interest (even those who do not attend). Or, a criminal justice system that protects the rights of the accused protects the dignity of anyone accused of a crime. This fact, in turn, explains why

our implementing and sustaining a particular kind of legal system is an activity that constitutes the good of each (even those who are not accused). Or, social welfare programs protect the well-being of the poor. This explains why our putting in place such programs is an activity that constitutes the good of each (even those who aren't poor).

All of this, I think, points toward a quite pervasive intertwining of interests. It suggests that many specific policies may actually promote everyone's good simultaneously. And *that* suggests that if some members are burdened in order to benefit the others—through the levying of a tax, or something like that—they can be compensated when benefits accrue to another. If so, then we might be able to justify policies that burden some in order to benefit others by appealing to compensation—by appealing to the fact that those burdened are left at least as well off as they were prior to the policy's implementation. As I have said before, this does not imply that just any burden may be permissibly imposed upon a person so long as she is compensated. But it does mean that some burdens that would have otherwise been impermissible to impose may become justifiable by virtue of the fact that the burdened individual is compensated. It means, that is, that we might have less to worry about, morally speaking, when the commonwealth imposes burdens on individuals in order to advance the overall good.

## **7. Answering Some Worries about the Very Idea of a Common Good**

The idea of a “common good constituting the good of the people”—along with the idea of a good in which everyone shares, and the idea of intertwining all citizens' interests—is one that 20<sup>th</sup> century liberals have tended to view with great suspicion, and so I would like to conclude this chapter by saying a few things about how my use of the concept here avoids some of the most common worries. There is, of course, one familiar kind of objector who maintains that *all* considerations about citizens' good or welfare fall outside of the state's legitimate sphere of concern—and I will consider such views in Chapter 6. More surprisingly, perhaps, is the slew of objectors who admit generally that citizens' good or welfare falls within the state's purview, but

who recoil at the thought of unleashing the state to promote something called the “common good,” or the “common interest,” or something along those lines. Such objectors find the idea of a common good troubling for various reasons. Sometimes the notion of a common good is thought to be hopelessly vague, and thus liable to be used as a mere pretext for imposing on a people a sinister political agenda.<sup>14</sup> Others insist that it could be no more than wishful thinking to believe that there is *anything* that could count as a shared good of all, and that by deceiving ourselves into thinking that such a thing exists, we risk failing to acknowledge the burdens inevitably borne by some individuals as a result of any political act.<sup>15</sup> In fact, it is not just philosophers who are common good-averse. Consider this definition of totalitarianism from an undergraduate-level American government text book: “Totalitarianism is the belief that government should have unlimited power. A totalitarian government controls all sectors of society: business, labor, education, religion, sports, the arts. . . **The object is to produce a perfect society serving some master plan for ‘the common good’**” (the bolding is mine, but the scare quotes around “the common good” appear in the original).<sup>16</sup> In the end, it appears that an innocuous-sounding concept like “the common good” has come to be closely associated with the darkest and most dystopian visions of totalitarianism.

So one concern was that the common good is indeterminate and vague, and that this invites political leaders (and philosophers) to simply *assert* that whatever policy they prefer constitutes the good of each. In this chapter, I have only given a general argument that society-wide group action is the right kind of thing to constitute an individual’s interest, and have said

---

<sup>14</sup> Joseph Schumpeter, for instance, tells us that “the common good” means different things to different people, and that its content is indeterminate. Aiming at such a vague abstraction, he believes, is a recipe for an irrationally ordered society—or perhaps something even worse. (See his *Capitalism, Socialism, and Democracy* (USA: George Allen and Unwin, 1976), pp. 250-256.)

<sup>15</sup> Isaiah Berlin warns us about this in *The Crooked Timber of Humanity*, ed. Henry Hardy (Princeton: Princeton University Press, 1998), p. 43. Berlin’s concern about the failure to register the inevitable burdens to individuals is pervasive in his work generally.

<sup>16</sup> Kenneth Janda, Jeffrey M. Berry, Jerry Goldman, *The Challenge of Democracy: Government in America*, 9<sup>th</sup> edition (Belmont, California: Wadsworth Publishing, 2006), p. 20.

nothing about how we would determine *which* actions constitute an individual's interest. So I may seem particularly vulnerable to the charge of advancing a view of the common good with no content—one that is just waiting for someone to arrive and insert their own dubious content. My intention, though, has not been to provide cover for anyone who wants to assert that her favorite vision of an ideal society really does constitute everyone's good. Rather, I hope to motivate philosophical inquiry into the question of which standards could be used to evaluate group activity and its contribution to an individual's living well. Indeed, I believe anyone should demand an argument before accepting the claim that something is a common good.

What kind of argument, then, other than bald assertion, could establish that one kind of group action is good and another is bad? What I imagine is that certain kinds of social criticism could be put to use in order to better understand why some activities are constitutive of—and others are destructive of—the common good. To give one example, Rousseau suggests that under certain inegalitarian societal arrangements, individuals seek to gain a certain kind of regard by dominating and degrading others. But when people look to those they degrade for regard, they simultaneously diminish the value of that very regard. After all, being regarded highly by a person who isn't worth treating well is arguably not regard worth having. Rousseau's account, then, could be used to critique the activity carried out by particular societies—the activity of sustaining inegalitarian societal arrangements. Such activity comes under scrutiny because it exhibits a deep internal tension, and, thus, the activity itself is revealed to be defective. This in turn suggests that each member would have her own interests advanced were the society to engage in sustaining institutions and societal arrangements of a different sort. Of course, this sketch is not a worked-out argument. But if my view that the common good is constituted by group action is correct, then I suspect that fruitful insights as to *which* activities constitute the common good might be found in social criticisms of this sort, and we might have new reasons for exploring the works of Rousseau, Hegel, Marx, and others.

The second worry was that aiming at the common good involves aiming at some good over and above the good of the members. This seems to be what Plato advocates in *The Republic*. Socrates tells us that within the ideal city, individuals share spouses, children, and property—all in the service of binding the city together, realizing *its* good.<sup>17</sup> When Adeimantus's expresses concern that the city's guardians will not be happy, Socrates retorts: "we aren't aiming to make [the guardians] outstandingly happy, but to make the city so."<sup>18</sup> Thus, it seems like Socrates is imagining a good of the city that floats free of the good of its members, capable of being realized without advancing the good of *any* citizen. If by "common good" we mean "good of the commonwealth" in this Platonic sense, then few will endorse the claim that we have *any* reason to promote it.

One could raise a version of Adeimantus's worry in response to my own view by saying: "On your view, society as a whole engages in activity that is good in some sense. But I only care about the individual. Who cares about the group living well, and why should we ever even consider giving such a thing any weight when deciding what to do?" There are two reasons why I think that my view is not actually open to this criticism. First, the view I have articulated is *not* that we should concern ourselves with the good of a whole that floats free of the good of individual persons. Rather my view is that what others do can constitute a good of mine (as well as of their own), and so what others do can constitute *my* living more or less well. Their success can be my success, and their failure can be my failure. Second, as I explained in the previous section, a group action itself is often valuable *because* it promotes the good of some individuals. Part of the reason one would claim that sustaining a health care system is in everyone's interest is that in sustaining such a system, health care is provided to some individuals who otherwise would not have it. So the view does, in this way, exhibit concern for individuals. So while the

---

<sup>17</sup> Plato, *The Republic*, trans. G.M.A Grube, revised by C.D.C. Reeve, in *Plato: Complete Works*, ed. John Cooper (Indianapolis: Hackett Publishing, 1997), p. 1089 (462<sup>a</sup>).

<sup>18</sup> *Ibid.*, 420<sup>b</sup>. See also 466<sup>a</sup>.



view that I have defended in this chapter seems similar to the kinds of views that many 20<sup>th</sup> century liberals worried about, there are specific features of my view that should assuage many of those worries.

## **8. Conclusion**

Let me conclude by simply restating the position for which have I argued in this chapter. The commonwealth is an agent, which is under a general normative pressure to perform good or worthwhile actions. But because doing so inevitably requires burdening some of its members, there is a worry about whether doing so is morally justifiable. One possible way to justify the burdens to those who are harmed is to compensate them. And I have shown that worthwhile collective action is, in itself, a common good, which provides compensation for each individual—compensation that may be used to justify burdening individuals. So, like an individual who simply realizes her good, knowing all the while that she is compensated for burdens that she must impose upon herself in doing so, a commonwealth can proceed knowing that all of its members share in the goodness of the worthwhile collective action in which is engages.

## Chapter 5: The General Will

### 1. Introduction

Thus far, I have argued that the commonwealth is a group agent, which realizes the common good by engaging in worthwhile endeavors. However, I haven't yet shown that an individual ought to prioritize the commonwealth's collective endeavors over her own. Despite the fact that an individual is under some normative pressure to cooperate with her political community, she still might, for all I have said, have most reason to pursue her own individual projects. So in this chapter, I will argue that a democratically constituted commonwealth has authority over its members, and thus that by making decisions democratically, a commonwealth makes it the case that its citizens ought to prioritize the collective endeavors that it chooses. In short, democratic institutions constitute a people's General Will.

Authority, however, is a puzzling concept. To say that X has authority over Y is to say that Y ought to follow X's directives, even if the directives prompt her to do something other than what the balance of reasons favors.<sup>1</sup> That an individual ought to do what the balance of reasons favors, however, seems like a fundamental requirement of rationality. Given this, it is difficult to see how genuine authority is even possible. So I will begin in sections 2 and 3 by critiquing a prominent account of authority, and then proposing my own account. Then, in sections 4-9, I argue that democracies do indeed possess authority.

### 2. The Service Conception

Attempting to illustrate the paradoxical nature of authority, Joseph Raz writes that, "to be subjected to authority, it is argued, is incompatible with reason, for reason requires that one should always act on the balance of reasons of which one is aware. It is of the nature of authority

---

<sup>1</sup> The word "authority" is sometimes used in another way, which simply implies that someone knows a lot about a topic or that someone has a lot of expertise. A person might, for instance, be an authority on Kant, or on saving for retirement. In fact, this sense of the word "authority" resembles in important ways the specific conception of authority that Joseph Raz offers (which I discuss in the next section).

that it requires submission even when one thinks that what is required is against reason. Therefore, submission to authority is irrational.”<sup>2</sup> The challenge, then, is to develop a conception of authority that does not render submission to it irrational.

The Service Conception is Raz’s influential attempt to meet this challenge.<sup>3</sup> In this section, I will spell out Raz’s view and raise a worry about it—a worry that will motivate me to propose an alternative conception of authority in the next section. The Service Conception is constituted by three theses. The first is that the directives of an authority are preemptory. This means that in an individual’s practical reasoning, an authority’s directive preempts, replaces, or excludes her own judgments about what she has most reason to do.<sup>4</sup> For instance, a restaurant employee might judge that the balance of reasons favors ordering more avocados. But if the manager possesses authority and directs her to make do with the avocados she has, then she simply ought to ignore what she believes the balance of reasons favors and comply with her employer’s directive.<sup>5</sup> This part of the Service Conception thus constitutes a denial of the paradox-generating claim that, “reason requires that one should always act on the balance of reasons of which one is aware.” A rational agent may sometimes act on reasons that preempt other reasons, rather than on the balance of reasons of which she is aware.

The second part of the Service Conception is what Raz calls the "Normal Justification Thesis." It says that, "the normal way to establish that a person [or corporate body] has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him. . . if he accepts the directives of the alleged authority as

---

<sup>2</sup> Joseph Raz, "Legitimate Authority," in *The Authority of Law: Essays on Law and Morality* (Oxford: Clarendon Press, 1979), pp. 3-27 (quotation from p. 3). Raz actually talks of the "paradoxes" of authority. The second pertains to autonomy, though I won't be discussing that here.

<sup>3</sup> Raz's earliest writings on the topic are contained in *The Authority of Law*. Raz lays out and argues for the Service Conception in "Authority and Justification," *Philosophy and Public Affairs* 14 (1985), pp. 3-29. His most comprehensive discussion of authority and of the Service Conception is found in *The Morality of Freedom* (Oxford: Clarendon Press, 1986), pp. 23-105.

<sup>4</sup> Raz, *The Morality of Freedom*, pp. 38-42, pp. 57-62.

<sup>5</sup> Importantly, the employer's directive is not just another reason to be added to the balance of reasons. The employee simply takes the employer's directive to preempt all the other reasons. (See *Ibid.*, p. 41.)

authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly."<sup>6</sup> For instance, suppose that each individual has reason to contribute to the common defense, and that a government consults experts and holds hearings to determine how best to defend the nation. It then directs its citizens to pay a tax in order to enact the plan it determines to be the best. The Normal Justification Thesis suggests that we can establish the government's authority to demand payment by arguing that the government's expertise enables it to make the best decisions about such matters, and thus that by paying the tax, the individual will make the contribution to the common defense that she has reason to. Beyond expertise, Raz thinks that being in a special position to solve collective action problems or to sustain worthwhile conventions might also serve to establish a person or corporate body's claim to authority.<sup>7</sup>

Third, the Service Conception holds that for a directive to have the force of authority, it must be, "based on reasons which already independently apply to the subjects of the directives and are relevant to their action in the circumstances covered by the directive."<sup>8</sup> For instance, an individual might have reason to help the poor, to contribute to the safety of the roads, or to invest her retirement money in a safe place. She has such reasons regardless of whether an authority exists. According to Raz, an authority must issue directives that are tethered to such reasons, and not to reasons that the subject has only by virtue of the authority's existence.

In a nutshell, an adherent of the Service Conception holds that sometimes an individual will best conform to the reasons she already has by allowing the directive of some person or entity to preempt the other reasons she judges herself to have. Thus, she sometimes ought to ignore the balance of reasons of which she is aware and comply with the directive. In such an

---

<sup>6</sup> Ibid., p. 53.

<sup>7</sup> Ibid., pp. 49-51.

<sup>8</sup> Ibid., p. 47.

instance, the issuer of the directive is a genuine authority, and it is rational for the subject of the directive to comply. "Authority" is thus rendered non-paradoxical.

According to the Service Conception, then, a democracy's authority, if it possesses it at all, does not derive from its granting each an equal voice per se. It derives from—and depends upon—its ability to issue directives that, when followed, enable its members to act for the reasons they already had.<sup>9</sup> A member who doubts the quality of a democratic body's directives should thus reject its authority, and thus deny that she ought to obey it. This, even if each member had an equal say in determining which directive was issued.<sup>10</sup>

There have been a number of important objections raised against the Service Conception. Here, I want to consider just one, which is that the Service Conception must either be based on an unsound theory of practical reasoning, or collapse back into a theory that directs the individual to act only upon the balance of reasons of which she is aware.<sup>11</sup> Uncontroversially, there are times when an individual, by adhering to a policy of obeying a particular commander's directives, will more often than not better comply with her reasons than were she to adhere to a policy of attempting to discern the reasons herself. But the question remains whether it is rational for such an individual to obey the directive of the commander *in each particular instance*. Obviously it is rational to comply in instances in which one has decisive reason to believe that the commander is correct about the reasons that one has. But in such instances, the balance of reasons of which one is aware simply favors doing whatever the commander directs

---

<sup>9</sup> David Estlund has recently argued that democracies tend to make good decisions, and that we should recognize their authority, in part, because of this. (See his *Democratic Authority: A Philosophical Framework* (Princeton: Princeton University Press, 2008).)

<sup>10</sup> In fact, many have objected to the Service Conception on grounds that the way that it conceives of authority rules out the possibility that authority depends upon procedure and process. For versions of this objection, see Thomas Christiano, "The Authority of Democracy," *Journal of Political Philosophy* 12 (2004), pp. 266-290; Jeremy Waldron, "Authority for Officials," in *Rights, Culture, and the Law: Themes From the Legal Philosophy of Joseph Raz*, eds. Lukas H. Meyer, Stanley L. Paulson, Thomas W. Pogge (Oxford: Oxford University Press, 2003), pp. 45-70; Scott Herschovitz, "The Role of Authority," *Philosopher's Imprint* 11 (2011).

<sup>11</sup> The objection mirrors almost perfectly a famous objection to rule utilitarianism, which says that rule utilitarianism either requires an individual to act irrationally or that it collapses back into act utilitarianism.

one to do. However, sometimes the subject will be aware of a strong reason against performing the commanded act. If the reason is *not* strong enough to outweigh the reason one has for believing that the commander is correct, then the balance of reasons still favors doing whatever the commander directed one to do. But if it is strong enough, then it is unclear how it could be rational to allow the directive to preempt the strong reason against carrying out the act, given that the likelihood of doing what one already had most reason to do is supposed to be what justifies obedience in the first place.

The dilemma, then, is this: if the individual ignores the strong reason against obeying, then she is irrational. If she does not ignore it, then she simply acts on the balance of reasons of which she is aware. So while Raz tries to use the notion of preemption to show that rationality does *not*, after all, require that an individual always act on the balance of reasons of which she is aware, it seems as if he has failed to do so. We are left, then, with obedience to authority either being irrational, or being indistinguishable from acting on the balance of reasons. In short, it doesn't seem that we have been taken beyond the puzzle with which we began, which is that an individual ought to do what the balance of reasons favors, and that doing otherwise is irrational—which leaves us without much room for genuine “authority.”

### 3. The Transparency Conception of Authority

The Service Conception's defense of obedience rests ultimately upon the normative force supplied by reasons that the individual already had, independent of the existence of a purported authority. This left us wondering why an individual would follow directives that she knows will lead her to act counter to those reasons. In this section, I will propose a conception of authority that I call the “Transparency Conception.” On the Transparency Conception, the defense of obedience does not ultimately rest on the normative force supplied by the reasons that the subject already has, independent of the existence of a purported authority. Instead, my account appeals to a different kind of normative force altogether—a kind of normative force generated

through what I call "transparency." On my account, an individual may, insofar as she is rational, act from the normative force generated through transparency instead of acting on the balance of reasons of which she is aware. So transparency plays the same role in my account that preemption plays on Raz's account. But the fact that there are two distinct types of normative force at work in my account, rather than just one, enables me to solve the puzzle about authority without falling prey to the objection that I leveled against Raz. I will begin, here, by distinguishing between the two kinds of normative force, and then conclude the section by laying out a conception of authority that rival Raz's.

To illustrate the distinction between the two kinds of normative force, let us begin by considering two kinds of mistakes that agents sometimes make. First, agents sometimes judge incorrectly which substantive reasons they have. By "substantive reasons" I simply mean features of one's situation (excluding intentions that she already has) that bear on a question of what should be done or intended. So, if you are deciding whether to buy a house, and you take the fact that your horoscope warns against making important life decisions today as a decisive reason not to buy, while ignoring the fact that you need a home and have the money, then you err by mis-judging your substantive reasons. If you commit this kind of error, I might try to place normative pressure on you to act differently by arguing against the reasons you took to be decisive—for instance, by arguing that your need for a house is a decisive reason.

Another kind of mistake, however, involves failing to act upon the substantive reasons that one judges to be decisive. For instance, you might judge there to be decisive reason to buy a home and to do all of the things that this entails, but then, because of laziness or fear, fail to contact the bank about securing a loan, fail to contact a realtor, fail to attend open houses, and so on. Such cases are familiar from discussions of *akrasia*. To get you to correct *this* mistake and to begin the process of buying a home, it would be futile to argue against the substantive reasons you took to be decisive. After all, you took there to be decisive reason to buy the home. So, *that* cannot be the problem. Instead I must point out that *you yourself* have decided to buy the home

on the basis of reasons that *you yourself* take to be decisive. I must point out to you that you yourself are not doing what you judge yourself to have decisive reason to do. Indeed, I might do this even if I disagree with you that the balance of substantive reasons favors buying a home—I might believe that you shouldn’t take there to be decisive reason to buy it, but that since you do, you ought to buy it.<sup>12</sup>

So irrespective of what substantive reasons there are to act or to refrain from acting, we generally think declarations like the following expose a rational error: “I take there to be decisive reason to be  $\phi$ -ing, but I’m not  $\phi$ -ing.” The agent’s specific error here, of course, consists in her behaving akratically. But I believe that this specific error belongs to a larger category of error, which involves running afoul of the normative demands of *transparency*.<sup>13</sup> (The term “transparency” is intended to highlight that the reason-taking is supposed to shine through to the acting.) In order to convince the reader that transparency constitutes a genuine and unique kind of normative force, I will now argue that the normative force of transparency cannot be reduced to the normative force of substantive reasons. Transparency cannot, I will argue, be reduced or explained away.

Imagine an account of agency that held that all practical normative force is ultimately derived from the substantive reasons that an agent has. Such an account might hold, for instance, that an individual not only has substantive reasons to act in particular ways, but that she has a substantive reason to do what she takes herself to have decisive reason to do—that she has a substantive reason to obey a rule against behaving akratically. That is, the account would say that since an agent is under normative pressure to be  $\phi$ -ing by virtue of the fact that she

---

<sup>12</sup> Thomas Aquinas famously believed that acting against one’s conscience is bad, even if one’s conscience is mistaken and one ultimately does what one had objective reason to do by going against it. (See *Summa Theologica* Parts 1-2, Question 19, articles 5-6.)

<sup>13</sup> Importantly, this has nothing to do with the way epistemologists and philosophers of mind sometimes use the word “transparency” when discussing the ability to know the content of one’s own beliefs by attending to the external world. Specifically, I warn against associating my use of “transparency” with Gareth Evans’s view in *Varieties of Reference* (Oxford: Oxford University Press, 1982), or Richard Moran’s view in *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press, 2001).



judges herself to have decisive reason to be  $\phi$ -ing, there must be a substantive reason that militates in favor of her  $\phi$ -ing when she judges that she should be  $\phi$ -ing. We sometimes make arguments of this form in order to defend inter-personal obedience. To give an example, my mother might claim that the fact that she has bestowed so many benefits upon me gives me decisive reason to do what she decides.<sup>14</sup> If I accept the argument, I might then say: “My mother takes there to be decisive reason for me to buy the home. *I take there to be decisive reason to obey the judgment of those who have bestowed great benefits upon me.* So I take there to be decisive reason to buy the home. Thus, I will buy the home.” The italicized sentence helps clarify this piece of practical reasoning by appealing to the putative fact that one has decisive reason to obey those to whom one owes a debt of gratitude. It is the normative force of this reason that we appeal to in order to explain why I ought to do what my mother decides.

Can we construct a similar argument for the claim that an individual is under normative pressure to do what she takes herself to have decisive reason to do? Such an argument might look like this: “I take there to be decisive reason for me to buy a home. *I take there to be decisive reason to do what I take there to be decisive reason to do (or, I have reason not to behave akratically).* So I take there to be decisive reason to buy a home. Thus, I am buying a home.” This is a problematic argument. Here the italicized sentence fails to clarify anything. Worse, by implying that a rational agent needs some additional substantive reason to do what she takes herself to have decisive reason to do, the inclusion of the sentence seems to suggest that the agent will require *another* substantive reason to act upon *that* reason. (Or, perhaps she will need a reason not to behave akratically with respect to her reason not to behave akratically). We are thus led quickly into an infinite regress. The lesson, then, is that an agent must eventually quit

---

<sup>14</sup> Samuel Pufendorf argues that we have such a reason, and uses it to argue that we should obey God. (See excerpt from *On the Duty of Man and Citizen*, in *Moral Philosophy from Montaigne to Kant*, ed. J. B. Schneewind (Cambridge, UK.: Cambridge University Press, 2003), p. 161.)

appealing to further substantive reasons, and simply act.<sup>15</sup> But since she commits a normative error if she fails to act on the reasons she took to be decisive, the normative force that she fails to respond to must be distinct from the force supplied by her substantive reasons.

So at this point, one has three options. (1) One can deny that the requirement to do what one takes oneself to have decisive reason to do is normative.<sup>16</sup> (2) One could insist that all normative requirements ultimately derive from what I have called “transparency.”<sup>17</sup> (3) One could claim that there are two distinct sorts of normative force—the force of transparency and the force of substantive reasons. So we must either deny altogether that an agent is under normative pressure to do what she takes herself to have decisive reason to do (option (1)), or we must acknowledge that transparency is a genuine type of normative force (options (2) and (3)). Niko Kolodny, in an influential paper, opts for the former. However, in the essay, he only briefly mentions the possibility of positing two distinct and irreducible kinds of normative requirements, and then seems to dismiss the possibility as unparsimonious.<sup>18</sup> Then, relying on the assumption that there are *not* two irreducible kinds of normativity, he proceeds to argue that the requirement to do what one takes oneself to have decisive reason to do is not normative. The surprising conclusion, then, is that contrary to appearances, there aren’t two kinds of normative mistakes that our agent looking to buy a home can make after all. I find this result too surprising. There is a long tradition, going back at least to Aquinas, that treats the dictates of one’s own conscience or mind as generating normative force. It seems to me that denying this in the name of parsimony is rather rash. So while I cannot delve too deeply, here, into Kolodny’s intricate view, I am going to proceed as if (2) and (3) are the most plausible options, and thus

---

<sup>15</sup> This is a point similar to the one made by Lewis Carroll in “What the Tortoise said to Achilles,” *Mind*, n.s., 4 (1895), pp. 278–280.

<sup>16</sup> See Niko Kolodny, “Why Be Rational?,” *Mind* 114 (2005), pp. 509–563.

<sup>17</sup> See Christine Korsgaard, *The Sources of Normativity* (Cambridge, UK: Cambridge University Press, 1996). Korsgaard believes that the normative force of substantive reasons ultimately derives from the fact that a rational agent takes some considerations to be reasons.

<sup>18</sup> Kolodny, “Why Be Rational?,” p. 512 (footnote 4).

will proceed under the assumption that transparency is a genuine kind of normative force (while admitting that pressing a view like Kolodny's would be a way of objecting to the view I lay out here).

So now I am in a position to lay out the Transparency Conception of authority. According to it, X has authority over Y by virtue of X's directive being transparent to Y's action. The conception solves the paradox with which we began by, like Raz, denying that reason requires acting upon the balance of reasons of which one is aware. However, since the normative force of transparency is not derived from the normative force of the reasons one already has (as it is for Raz), the Transparency Conception does not face the dilemma that Raz did. A rational individual can respond to the normative force of transparency despite knowing that, in so doing, she will act against her reasons. There is no puzzle about this because the normative force of transparency does not derive from—nor does it purport to derive from—the agent's substantive reasons.

According to the Transparency Conception, an individual has authority over herself by virtue of her judgment's transparency to her action. Of course, the individual will normally experience no conflict between the normative force generated by the balance of reasons of which she is aware and that generated by the transparency of her practical judgment. But that is because her practical judgment is usually based on the balance of reasons as she sees them. More interesting and controversial, though, will be cases in which X and Y are not identical. To give an example relevant to the present topic, a political body could have authority over the members of a community or commonwealth. For instance, we might claim that a democratic body constitutes a people's General Will, which makes decisions whose normative force shines through to the citizens' action. If so, then we might think that the following thought is cogent and in no need of normative support from further substantive reasons for obedience: "The democratic body judged there to be decisive reason for me to be  $\phi$ -ing, and so I am  $\phi$ -ing." Were democracies to possess genuine authority, the Transparency Conception would say that pointing

out that the democratic body decided that I should  $\phi$  is not merely to draw my attention to a reason like any other. It is more like drawing my attention to the fact that *I* have decided to do something on the basis of reasons that I took to be decisive. While we wouldn't call the disobedient citizen "akratic," she and the akratic agent have something in common. They both err by ignoring the normative force generated through transparency.

#### 4. Establishing Authority

One thing we might reasonably demand from a conception of authority is some sense of how we could establish, through argument, that an individual or corporate body possesses it. But it is natural to worry that the Transparency Conception cannot meet this demand. After all, one might assume that the only way to argue that an individual ought to obey a directive is to appeal to substantive reasons that bear on the question of whether to obey. However, this would mean that the transparent normative force would ultimately derive from substantive reasons, and as I argued in the previous section, that cannot be the case. So this strategy is not a promising one. Another strategy is intuitionistic. That is, one might claim that it is a mistake to attempt to prove that someone has authority, and that the fact that someone's directives are transparent can only be apprehended directly and immediately through intuition.<sup>19</sup> However, many find appeals to direct apprehension either suspicious or disappointing. So I want to show that adherents of the Transparency Conception could actually argue, or attempt to prove, that someone or some thing possesses authority. To do so, I will employ an argumentative strategy developed by Thomas Nagel.

In *The Possibility of Altruism*, Thomas Nagel contrasts two argumentative methods, which he calls "interpretation" and "justification."<sup>20</sup> When we use justification to argue in favor

---

<sup>19</sup> I am imagining a view similar to H. A. Prichard's view about moral truths, which he thinks cannot be proven by non-moral argument, but must be apprehended directly. (See his "Does Moral Philosophy Rest on a Mistake?" *Mind* 21 (1912), pp. 21-37.)

<sup>20</sup> Thomas Nagel, *The Possibility of Altruism* (Princeton: Princeton University Press, 1970), pp. 18-23.

of acting in a particular way, we, "proceed within the context of a system of reasons, by showing that certain conditions are met which provide sufficient reason for that [action] which is being justified."<sup>21</sup> For instance, you might claim that an individual ought to aid the poor, and then point out that obeying the tax law is a way of doing this efficiently.<sup>22</sup> This would be a way of justifying obedience to my government. Interpretation, on the other hand, "is an attempt to link . . . practical principles to. . . basic features of the conception which each person has of himself and of his relation to the world, and to link the two in such a way that adherence to the principles can be seen as a practical expression of the conception."<sup>23</sup> This is very abstract, but the crux of the idea is that providing an interpretation involves arguing that an individual's self-conception demands that she acknowledge the normative force of certain principles. Nagel uses this method of argumentation to demonstrate that an individual is normatively bound by certain practical principles. The principle of prudence directs an individual to treat substantive reasons that she will have in the future as normatively significant in the present. For instance, it directs a person who will have reason to know Italian when she travels to Italy next month to learn some Italian *now*. Nagel gives several arguments demonstrating that for an individual to reject the principle of prudence—for her, for instance, to fail to treat the fact that she will need to speak Italian when she arrives in Italy next month as a reason to learn some Italian now—is for her to deny, by implication, that she has a future of her own. He then claims that a person conceives of herself as a temporally extended being, with a future of her own. Conceiving of herself in this way, Nagel says, is "inescapable." Because of this, she cannot coherently reject the principle of

---

<sup>21</sup> Ibid., p. 18.

<sup>22</sup> This highlights that the Normal Justification Thesis is a form of "justification" in Nagel's sense.

<sup>23</sup> Nagel, *The Possibility of Altruism*, p. 18.

prudence.<sup>24</sup> So, a person's adherence to the principle of prudence is, Nagel thinks, a practical expression of her self-conception.<sup>25</sup>

Now, Nagel insists that an individual does not have a "reason" to conform to the principle of prudence.<sup>26</sup> We do not establish the normative force of this principle by appealing to some *other* considerations that militate in its favor, but by showing that recognizing its normative force is an expression of a self-conception that we find inescapable. So it seems that Nagel establishes the normativity of the principle of prudence without appealing to the normative force of reasons. In "interpretation," then, I believe we have an argumentative method with the potential to establish that a person or corporate body possesses authority—a method that does not simply appeal to substantive reasons that militate in favor of obedience. In the case of a political body, for instance, we might argue like this: X is a particular kind of government. Y is a subject. X has authority over Y if Y's self-conception demands that she recognize the transparency of X's directives. This, in essence, is how I will go on to argue for democracy's authority.

### 5. Political Judgment and Collective Activity

Using the method of interpretation, I will argue that treating democratically enacted directives as transparent to one's action is a necessary practical expression of an individual's self-conception as an agent. Taking oneself to be an agent requires recognizing a normative relationship between one's practical judgment and what one does—indeed, she must recognize her judgment as transparent. Agents decide what to do on the basis of reasons, and if an individual does not take her practical judgment to put her under pressure to act, she attempts to dissociate herself from her judgment (or from her action) in a way inconsistent with thinking of

---

<sup>24</sup> Ibid., pp. 33-76.

<sup>25</sup> Ibid., pp. 99-124.

<sup>26</sup> Ibid., p. 18.

herself as an agent. My argument will ultimately be that denying democracy's authority implies a rejection of this self-conception. The argument hinges on a claim about how an agent must view her relationship to particular kinds of judgments she makes and particular kinds of activities she engages in. So before I argue for this claim, I must say a bit about these particular judgments and actions.

I will begin with activity. We often attribute certain happenings in the world to collections of people in a very loose sense—"attribute" in a way that does not necessarily imply that the group carried out fully-fledged action. So we might say of a particular multitude (S), that, "*S did A*," or, "*S contributed to A*," or, "*S failed to A*." Such attributions are sometimes made when the members of S bring about a state of affairs with their individual actions in aggregate, and sometimes they are made when the people all organize themselves so as to bring about a state of affairs intentionally. The point is simply that these attributions assign causal responsibility to the group for some happening for which no individual alone could be causally responsible.

Most individuals belong to various collective bodies whose collective behavior, broadly construed, is responsible for the existence of distinct societies and public spaces, each with their own distinctive features. It will be easiest to clarify this claim through examples. A person, along with those in her relatively immediate vicinity, often maintains a public space that we normally call a city or municipality. Whether snow renders the city's sidewalks impassible and whether its parks are filled with litter are not "up to" any individual resident. These features of the public space are determined by whether residents in aggregate tend to litter or shovel the walkways outside of their homes, or by whether they instead address these issues collectively by organizing garbage pick-up or snow-plowing. To give another example, a person, along with those in her less immediate vicinity, often maintains a society that we normally call a nation. As John Searle has argued, what counts as money and what counts as a legitimate exchange of goods within a region depends upon the collective recognition of a people—often of the people

that comprise a nation.<sup>27</sup> For a piece of paper to count as “money,” the nation must *do something* to make this the case—it must collectively decide to treat the paper as if it has value. To give a final example, a person, along with all of the others in the world, maintains a society (of sorts) that we sometimes call the “international community.” Whether there exists an extensive de facto right to move about freely, or whether there are tightly controlled borders, is not determined by any one individual, but by the international community, its collective recognition of such rights and boundaries, and the respect that members in aggregate afford the rights and the borders.<sup>28</sup> In each of these examples, it is either the individuals’ behavior in aggregate, or organized collective activity—not the will of any one person—that explains why the society or public space has the features it does. This is why we say that, “The city does a good job removing snow,” or, “The nation decided to recognize the Euro as its official currency.”

Proceeding now to the topic of judgment, there is a wide range of judgments that an individual could sensibly make pertaining to the collective body’s behavior and the states of affairs it realizes. For instance, one might judge that, “*For S to A would be good*,” or, “*For S to avoid doing A would be just*,” or, “*It would be most efficient for S to A*.” More concretely, a public space’s physical condition could render it more or less convenient to traverse; some economic practices are more or less just; some immigration policies are more or less humane. One can also pass judgment on the *way* that a collective body realizes and maintains such features—judging that, for instance, “*The best way for S to A is to do B*.” Most importantly, one can make second-order judgments about the normative significance of these other judgments—judgments such as: “*Because for S to A would be just, S must A*.” So, for instance, one could judge an economic practice to be unjust, and then make the second-order judgment that the

---

<sup>27</sup> John Searle, *The Construction of Social Reality* (New York: The Free Press, 1995), p. 35, pp. 37-42.

<sup>28</sup> Searle discusses rights on *ibid.*, p. 93. See also his *Making the Social World: The Structure of Human Civilization* (Oxford: Oxford University Press, 2010), pp. 174-198.



injustice is sufficient grounds on which to justify abolishing the practice, regardless of how efficient it is. I will call all of the aforementioned kinds of judgments “political judgments.”

Now, do we have any reason to suspect that making political judgments about one’s own society carries any normative significance for an agent? I think we do. After all, when I make a political judgment about my own society, it isn’t merely an academic matter, as it would be were I to make a judgment about an ancient or distant society. When I make a political judgment about my own society, it is a judgment about what kind of society *I*, with other members of the collective body, should be sustaining. Normally when an agent makes a judgment that is relevant to her own action, the judgment has direct implications for what she will do next, insofar as she is rational. Identifying what, exactly, is required of an agent, given her judgment, is not always a simple task. *Something*, however, is almost always required. In the most familiar case, an agent who judges that she ought to undertake a course of action is required to undertake that very course of action. Or, if this is impossible, perhaps she is required to put herself into a position where it *would* be possible. Or, if even that is impossible, perhaps expressions of regret are in order.

The purpose of the above discussion has been to motivate this question: as an agent, what is an individual normatively compelled to do when she makes a political judgment about things done by her city, her nation, or the international community? Again, it is a pressing question, because the judgment is the agent’s own, it concerns matters relevant to *her own action*, and an agent’s practical judgment is transparent to her action. In the following section, I will consider why one way of answering this question is unacceptable, with the hope that this will provide a clue as to what answer would be acceptable.

## **6. Taking One’s Own Political Judgment to Be Decisive**

In this section, I will consider the philosophical anarchist’s answer to the question raised above, which is that insofar as she is rational, an agent must take her own political judgment to be

decisive.<sup>29</sup> More precisely, on this view a member of a collective body carries out a deliberative procedure in her own mind that aims to identify which features her society ought to possess, and also to determine how these features ought ideally to be realized. She might make judgments about which rights should be recognized, or about whether there should be restrictions on immigration, for instance. Then she takes herself to be normatively compelled to advocate for her view and to cooperate *only* with those collective behaviors consistent with that judgment. Robert Paul Wolff endorses something in the vicinity of this position. Wolff argues that an agent has an overriding duty to act in ways that she judges to be good or wise, and that recognizing the authority of the law runs afoul of this duty, since recognizing the authority of the law requires setting aside one's judgment and participating in collective efforts that one does not endorse. For this reason, Wolff says that, "Insofar as a man fulfills his obligation to make himself the author of his decisions, he will resist the state's claim to have authority over him."<sup>30</sup> He thus embraces the anarchic conclusion that an agent's own judgment is decisive, leaving the law normatively impotent. In this section, I will advance an argument against this anarchist view. I will do this not simply for the purpose of demonstrating that the anarchist is mistaken, but because I believe that doing so will ultimately help us see why democracy possesses authority.

Even if one does not agree with Wolff's anarchistic view, most agree that *something* goes wrong with an agent who simply complies with directives that she disagrees with, without further justification for doing so. But it will help, here, to consider a bit more deeply why this is so. In the *Discourse on Inequality*, Rousseau takes up this very issue. He begins by contrasting an agent's relationship to her own action with her relationship to a piece of property, writing

---

<sup>29</sup> Philosophical anarchism is not the view that there should be no government, but instead the view that no governments possess genuine authority.

<sup>30</sup> Robert Paul Wolff, *In Defense of Anarchism* (Los Angeles: University of California Press, 1998), p. 18.

that, “[G]oods I alienate become something altogether foreign to me.”<sup>31</sup> I can transfer a piece of property to another party, and in so doing dissociate myself from it in such a way that I no longer have a right to it or responsibility with respect to it. Thus, according to Rousseau, the abuse of such goods can become, “a matter of indifference to me.”<sup>32</sup> But an individual’s action is unlike property in this regard, in that an agent cannot distance herself from her action in such a way as to give a third party complete responsibility for what she does. According to Rousseau, when I act, I always risk, “incurring the guilt of the evil I . . . commit.”<sup>33</sup> This becomes evident when we confront an agent who obeys a command directing her to do something bad or evil, and she responds that she was “just following orders.” Rousseau’s point is that insofar as we treat the individual as an agent with a will of her own, we won’t find such a reply to be sufficient. We are committed to the idea that an agent *moves herself* and does so *for reasons that she recognizes*. So when an agent acts, we take her, as the one who took there to be a reason, to be answerable for what she does. This is just part of what it is to be an agent. To simply obey a command that runs counter to the reasons of which one is aware is to be moved by what *someone else takes to be a reason*. It is to be moved by someone else’s judgment. And so were an agent simply to do whatever she was directed to do, she would dissociate herself from her own action—she would view someone else as the author of what she does. But conceiving of oneself as an agent rules this out—conceiving of oneself as an agent requires viewing oneself as the author of one’s actions. This, at least, is what I think Rousseau is trying to say, and it is for this reason that he thinks it appears problematic to claim that the law has authority.

---

<sup>31</sup> Jean-Jacques Rousseau, *Discourse on the Origin and Foundations of Inequality Among Men*, ed., trans. V. Gourevitch, in *The Discourses and Other Early Political Writings* (Cambridge, UK: Cambridge University Press, 1997), p. 179.

<sup>32</sup> Ibid.

<sup>33</sup> Ibid.

Rousseau's argument might cause us to question the whole notion of political authority, pushing us toward an anarchistic position like Wolff's.<sup>34</sup> I will argue, though, that this is not the conclusion we ought to draw. I will show that by taking her own political judgment to be decisive in the way that anarchism requires, a citizen actually saps her judgment of its normative force. Or to put the point paradoxically, we might say that by associating herself so strongly with her political judgment, a citizen actually dissociates herself from it. It will clarify my argumentative strategy a bit to give an analogy. Imagine a woman who believes that women are defective as thinkers, and thus are not warranted in holding epistemic convictions. Besides having a belief with dubious content, this anti-feminist has a more formal problem, which is that she has adopted an attitude that undermines *that very same attitude*. After all, if women are not warranted in holding beliefs, then the anti-feminist lacks warrant to believe that women are defective as thinkers. The paradoxical result is that the more strongly she avows her anti-feminist belief, the more strongly she dissociates herself from it. I will argue that an analogous paradoxical result is generated when someone treats her own political judgment as if it exerts overriding normative force upon her.

To show this, I will first make a point about the inevitability of political disagreement in civil society, and then a point about the proper attitude to take toward those with whom we disagree. Once we have these two points in view, it will become clearer how taking one's political judgment to be decisive is problematic. The first point is that civil society is constituted by a plurality of rational individuals whose political judgments conflict. By calling individuals "rational," I mean that each individual (who isn't mad, a small child, intoxicated, and so on) makes judgments that are transparent to her action.<sup>35</sup> Now, the existence of a plurality of rational individuals always carries with it the possibility of disagreement. In fact, we usually take it for granted that significant disagreement *inevitably* exists within many collective bodies—

---

<sup>34</sup> Though, of course, Rousseau does not draw this conclusion.

<sup>35</sup> I don't mean to suggest, however, that they are equally competent judges of what there is reason to do.

especially those collective bodies that interest me here. Whether those bodies are small communities whose collective behavior explains the conditions of the roads, or entire nations whose collective behavior explains the shape of various economic practices, disagreement is usually thought to be pervasive and unavoidable.<sup>36</sup> This disagreement is not likely to be resolved through appeal to experts—that is to say, through appeal to Razian authorities who have better epistemic access to the reasons than the members have. After all, within any populace there are serious disagreements about who the genuine experts are. American society, for instance, is deeply divided on the question of who possesses true economic expertise. But also, recall that our political judgments concern not just technical issues, such as which economic policy would create more jobs. They concern values—they are about whether it matters more to protect property rights or to provide health care, whether security or civil liberties are more important, whether equality or efficiency is better, and so on. It would be dubious to suggest that there would be widespread agreement on who would count as the experts about these issues.<sup>37</sup> Furthermore, disagreement is unlikely to be resolved by appealing to substantive reasons to obey the majority. While it might be plausible to argue, for instance, that a citizen ought to obey the democratic body out of respect for others' interest in having a say in determining her society's law,<sup>38</sup> there will be widespread disagreement about how serious a reason this is, and about whether other important reasons outweigh this one. So, while societies where everyone agrees are not metaphysically impossible, we generally hold that consensus is a practical

---

<sup>36</sup> There are many potential explanations for the inevitability of such disagreement. For instance, it might be that reason itself doesn't deliver a single unique correct judgment, at least with respect to many questions. Or it might just be that political questions are so complicated that finite and imperfect humans considering those questions are condemned to arrive at different answers. For a famous discussion relevant to this, see John Rawls on the "burdens of judgment" in *Political Liberalism, expanded paper edition* (New York: Columbia University Press, 1996), pp. 54-58.

<sup>37</sup> John Stuart Mill, however, tries to argue that such agreement is possible. (See his "Representative Government" in *Three Essays: On Liberty, Representative Government, The Subjection of Women* (Oxford: Oxford University Press, 1975), p. 282.)

<sup>38</sup> Thomas Christiano, for instance, argues that democratic decision-making affords each person an equal and fair opportunity to affect public policy, and that this fact gives citizens a reason to recognize the authority of democratic institutions. See his "The Authority of Democracy."

impossibility for all but the smallest, most homogeneous collectivities. So when arguing for claims in political philosophy, it should not be terribly controversial to appeal to the fact that disagreement exists.

My second point is about the attitude a citizen should take toward other agents and their political judgments. We have already concluded, with Rousseau, that an agent must, insofar as she is an agent at all, treat her own political judgment as exerting normative force upon her. If a citizen acknowledges that *she* must treat *her own* political judgment this way, it seems clear that she ought also to acknowledge that *others* must treat *their own* judgment as exerting normative force on *them* as well. After all, insofar as one recognizes that one's fellow citizens are agents, an individual must recognize that they are under the same sorts of formal normative pressures that she, as an agent, is. So even when one of my fellow citizens makes a judgment with which I disagree, I can acknowledge that she is under normative pressure to act in ways that accord with her judgment. And now extending this point to the anarchist citizen who takes her judgments to be decisive, if she believes that *she* ought to take *her* judgments to exert overriding normative force upon her, then certainly she ought to acknowledge that *others* ought to take *their* judgments to exert overriding normative force upon them as well.

Here, then, is the key move in the argument. If we combine the point made above with the point about the inevitability of disagreement, we see that a citizen who treats her political judgment as decisive should acknowledge that she is under overriding normative pressure to do things that inevitably conflict with things that her fellow citizens are themselves under overriding normative pressure to do. That is, she will judge, "I am under overriding normative pressure to be contributing to our  $\phi$ -ing," and that, "Others are under overriding normative pressure not to cooperate with our  $\phi$ -ing, and to oppose it instead." Recall, however, that political judgments concern *collectively* realized societies and public spaces. That is, the only way for the world to conform to one's political judgment is through collective activity. So, given what we have said about disagreement, the anarchist agent must judge that, "Insofar as

everyone is rational, we will not be  $\phi$ -ing.” Thus, in the face of disagreement among people who each ought to take her own judgment’s normative force to be overriding, there is a sense in which an agent must also acknowledge that her own political judgment lacks relevance with respect to what should *actually get done*. That is, our citizen should admit that the societal features that she judges ought to be realized, in fact, will not be realized insofar as those who are needed in order to realize them act rationally—which implies that it is not the case that such features ought to be realized after all.

I should pause here to clarify. One might object that there is a sense in which the citizen *could* maintain that the features she judges best ought to be realized. Namely, she could maintain that if everyone was sensitive to the right reasons, everyone would agree on which features should be realized, and all could participate in realizing those features without reservation. If this is right, then it might also seem right to say that what others ought really to do is make better judgments about the reasons they have, and act upon those reasons.<sup>39</sup> Recall, however, that we are taking disagreement to be a fact of life in civil society. In so doing, we treat a collective body that constitutes a city or a nation to be different from a small group, like a jury, within which consensus on the basis of right reasons remains within the bounds of practical possibility. So we can’t say that the all-things-considered thing to do is for all citizens to agree on the right reasons, since we have ruled out this possibility.<sup>40</sup> And once we do this, we get this

---

<sup>39</sup> Would this imply, then, that there exists no rational requirement to cooperate only with projects that one thinks there is decisive reason to carry out, and that the only requirement is to act for the right reasons? Not necessarily. The requirement to cooperate only with projects that one thinks there is decisive reason to carry out simply needs to be given what is sometimes called a “wide-scope” reading. The requirement is thus rendered: “You ought (either judge that there is decisive reason to carry out project X, or refuse to cooperate with project X).” An agent can meet the requirement in one of two ways: she can refuse to cooperate or she can alter her judgment about what there is decisive reason to do. The objection, then, could be that when I notice that others disagree with me, I should not think that they ought to meet the requirement by refusing to cooperate. Rather, they ought to meet the requirement by altering their judgment.

<sup>40</sup> Even if one is inclined to give the requirement the wide-scope reading mentioned in the previous footnote, once we eliminate the possibility of everyone meeting the requirement by judging the reasons correctly, we see that the only way for everyone to meet the requirement is by cooperating only with projects that they believe there is decisive reason to carry out.

result: our citizen holds that certain societal features ought to be realized, and simultaneously commits herself by implication to the claim that it isn't the case that those features ought to be realized (again, because insofar as they are rational, the people who are needed to do the realizing are under overriding normative pressure not to cooperate). This isn't a formal contradiction. But even if it does not trouble the logician, the result ought to trouble the citizen. The more she insists that her own judgment has absolute authority, the more the citizen undercuts her ability to truly commit to the claim that the societal features she judges should be put in place should actually be put in place after all—a paradoxical result, akin to the paradoxical result we witnessed in the anti-feminist case.<sup>41</sup>

To illustrate this argument with one of our stock examples, we might imagine a nation where there is disagreement about which economic practices to adopt. Each member takes her own judgment to be normatively overriding, and so recognizes that each other member ought to take her own judgment to be normatively overriding as well. Each will, that is, only comply with tax and financial regulatory policy that conforms to her own judgment, and will recognize that others ought only to comply with policy that they agree with as well. Since an economy's character is the result of an entire society's activity, and each is under overriding normative pressure to do something different from the others, each should recognize that there is a sense in which it is not the case that her preferred economic system ought to be put in place. But to make a practical judgment—to claim it as one's own—is usually to get behind it in a particular way. It is to commit to the idea that something should actually happen going forward, and that I, as the agent making the judgment, should play a role in what happens. But when I imply that it isn't the case that the economic practice that I judge ought to be implemented actually should be implemented, I undermine that commitment. I make the judgment, but simultaneously distance myself from it.

---

<sup>41</sup> The analogy is not exact, though. In the anti-feminist case, it was the content of a belief that implied dissociation between the woman and her belief. In my argument, it will be an agent's attitude toward the status of her political judgment that implies the dissociation.



Rousseau's point—the one we suspected might help to justify philosophical anarchism—was that one's self-conception as an agent precludes viewing oneself as dissociated from one's judgment and action. Now we have seen that affirming the anarchist position implies dissociation of this very sort. So anarchism is, perhaps surprisingly, ruled out insofar as one conceives of oneself as an agent. What is needed, I think, is a position somewhere in the middle between treating one's own political judgment as supremely authoritative, and treating it as if it lacks any authority at all. This is what I will propose in the following section.

### 7. Equal Voice and Democracy

Each agent must acknowledge that her own political judgment exerts some normative force upon her—and this is, as I argued, the normative force of transparency. Each must also recognize that each other agent is in the same position, under the same pressure from her own political judgment. The symmetry between each agent's situation, along with the fact that each ought to recognize this symmetry and also recognize that disagreement is inevitable, suggests the following hypothesis: an agent ought to treat her own political judgment as if it exerts some transparent normative force upon her, but no more and no less normative force than is supplied by the political judgments of each other member of a collective body to which she belongs. More colloquially, an agent should recognize the equal voice of each. In this section, I will fill out and defend this hypothesis by taking up three tasks. First, I will explain what I mean when I say that an agent ought to recognize the equal voice of each. Second, I will explain why agents who recognize the equal voice of each avoid implying the kind of dissociation implied by the anarchist position. Third, I will distinguish the view that I advocate from a natural misunderstanding of it.

I will begin with the issue of *how*, exactly, an agent treats her political judgment as exerting no more and no less normative force on her than the judgment of each other member of a collective body. My view is that this is possible if a collective body has decision procedures in

place that give each an equal voice. Whether the collective body is constituted by a small number of people whose activity is responsible for the condition of the roads, or a world-wide body whose activity is responsible for the existence of certain human rights, the body is *democratic* if it makes decisions through such procedures. A member then can acknowledge the equal voice of each by treating its decisions as transparent to her action. So just as a rational individual sometimes carries out a decision procedure in her own mind—weighing reasons and so forth—and then treats the decision to be transparent to activity within her private life, I am claiming that a citizen, insofar as she is rational, will treat a democratic procedure carried out *outside of her mind* to be transparent to her activity within the public sphere.<sup>42</sup> To return to the notion of the General Will, the idea is that a collective body could, through democratic procedures, decide to do something, and that individuals would cooperate insofar as they are rational. Just as a rational individual doesn't need a substantive reason to do as her own individual will decides, neither does she need a substantive reason to do as the General Will decides. The General Will, like the individual will, ought to be treated as *transparent* to an agent's activity.

Giving a full account of the form that institutions would have to take in order to grant equal voice, and thus to constitute the General Will, is not something I can do here. But very roughly, it seems reasonable to begin with a baseline on which everyone is allotted a single, equally weighted vote that she may use to register her political judgment. It also seems reasonable to give each an equal opportunity to place items on the agenda—to help determine what is discussed, considered, and voted on. Furthermore, it seems reasonable to strive to give each individual who has an opinion a forum to make her case, or at least a representative to whom she may present her reasons. From this baseline, we might make adjustments either in the name of equal voice, or to satisfy other desiderata (so long as equal voice is preserved). For

---

<sup>42</sup> I am working here with an intuitive distinction between the public sphere—which we constitute together—and the private sphere. More work would be needed to draw a sharp line between the two (or to argue that it can be drawn). This would be necessary in order to determine whether the General Will or one's individual will has authority over some particular set of issues.

instance, many issues about inequality of wealth and the ability of the rich to exercise disproportionate influence over deliberations need to be addressed if the process is truly to give each an equal voice. Another issue: some societies have persistent minorities. A procedure in which decisions are made by a simple majority vote, it might be argued, would not actually give such minorities equal voice. For such societies, then, we might need to make some adjustments from our initial simple baseline of “one person one vote.” There are also questions about how voting procedures could be designed in such a way as to avoid creating various inconsistencies or paradoxes that can be generated when trying to aggregate individual judgments.<sup>43</sup> But my purpose here is not to address these specific questions of institutional design. Rather, my purpose is to argue more abstractly that when a collective body has in place institutions that live up to certain democratic standards or live up to certain democratic ideals, one ought to obey laws enacted by those institutions. The specifics of how, exactly, to design institutions that live up to these standards and ideals, and that avoid various problems and paradoxes identified by social choice theorists, is a topic for another day.

My second task is to explain why an agent’s granting each other person an equal voice allows her to avoid the dissociation problem. The reason is that acknowledging the equal voice of each reinforces the normative force of one’s own judgment, even as it fails to grant it absolute authority. If a citizen believes that she is compelled to treat each other’s political judgments as normatively significant, then she ought also to believe that others are compelled to treat *her* political judgments as normatively significant.<sup>44</sup> So when she registers her judgment, she takes

---

<sup>43</sup> For a recent discussion, see Christian List and Philip Pettit, *Group Agency: the Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011), esp. chapter 2. They suggest several ways that procedures could be constructed to avoid the paradox they identify, some of which seem to me more plausible candidates for meeting the standard of equal voice than others.

<sup>44</sup> My point is that one affirms the status of one’s own judgment by recognizing the normative status of others’ judgments. Rousseau makes an analogous point about interests and about moral standing generally. For Rousseau, the General Will is a substantive notion (as opposed to a procedural one, which is how I have understood it). To act from the General Will is to recognize the interests and the moral standing of others, and by acting from the General Will Rousseau thinks the individual affirms her own status as a being with moral standing whose interests deserve consideration from others. Nagel makes a

the judgment to have normative significance for each other member—for all those who are needed in order to realize a society or public space that conforms to the judgment. The citizen's political judgment is thus normatively *empowered* for the same reason it was disempowered when she took it to be decisive or to exert overriding normative force. To spell this out a bit more precisely: (1) Each person occupies a perspective from which she makes a political judgment. (2) Each recognizes that her own political judgment exerts some normative force upon her. (3) Each recognizes that each other person also occupies a point of view from which political judgments are made. (4) Each recognizes that each other person is under the same pressure to recognize the normative force of her own judgment. (5) Each recognizes the equal normative force of each other person's judgment, and thus maintains that others ought to recognize the normative force of her own judgment as well. (6) By registering a political judgment with her vote, each thus places normative pressure on those who are needed to realize a public sphere that conforms to the judgment. Thus, an agent can make a political judgment and hold that it has implications for what should *actually get done* in civil society.

My final task is to forestall a natural misreading of my view. I am not claiming that a person recognizes the normative force of others' judgments *in order to avoid dissociation*. Putting it this way makes it sound as if I believe that agents have a substantive reason to try to avoid dissociation, and that they ought to obey democratic decisions as a way of acting for that reason. Instead, what I hoped to do was show that certain preconditions for conceiving of oneself as an agent entail that a person ought to recognize the equal voice of each—at least within a civil society. To revisit our analogy, in my anti-feminist example, the argument was not that the woman should feel compelled to draw inferences from her belief *as a means* to avoiding epistemic dissociation, as if she had some reason for not wanting to be dissociated. The claim was that treating herself as an epistemic agent at all required treating her own beliefs as having

---

similar point about reasons, arguing that a person affirms her own status as a "real" person with reasons of her own by treating the reasons of others as normative for her. (See *The Possibility of Altruism*, Part III.)

epistemic normative force. I intend my own account to be taken in a similar spirit. It isn't that we recognize the equal voice of each *as a means* to avoiding political dissociation. Recognizing the equal voice of each within civil society is demanded by one's self-conception as an agent, and we see this most vividly when we notice the dissociation that results from rejecting the demand.

## 8. A Possible Dilemma

Any actual decision procedure is capable of producing morally repugnant decisions. Democratic procedures are no different. Most think that a citizen has a duty to disobey certain laws, even if the procedure through which the law was made is legitimate. My view, however, might seem to lack the resources to allow for such disobedience. After all, the deliverances of the General Will are supposed to be transparent to action—they are supposed to determine what the citizen does insofar as she is rational, and she is not supposed to weigh reasons for and against obeying. So it seems like my view faces a dilemma. On one horn of the dilemma, an individual is required to cooperate with morally atrocious policies so long as they are chosen through democratic procedures. Or, on the other horn, the individual would assess the moral implications of the democratic body's decisions and decide for herself whether to cooperate. But saying this would take a lot of the bite out of my account, as the decision of the democratic body would then seem not to be so transparent after all.

Though I will not do it here, I believe that the way to avoid this dilemma is to develop a theory about the sorts of considerations that can *block* transparency. To see what I mean by this, consider first an analogy with individual judgment. As I have argued, there is transparency between an individual's taking there to be decisive reason to do something and her activity. It is possible, however, for an individual to engage in the reasoning process by which she takes a consideration to be decisive—a process that is normally transparent to her activity—but then identify conditions that undermine the process's authority. To give a quick example: an agent might, through a normally transparent reasoning process, decide upon a course of action that

seems straightforwardly inconsistent. Maybe it requires being in two places at the same time, or maybe it commits every dollar she has to two separate projects, and she only realizes this after she has assessed her reasons and adopted the plan. In such a case, a rational agent will have to abort the plan, or go back and re-evaluate. Transparency is blocked. The important point here is that such cases needn't cast doubt upon a thought process's transparency under normal circumstances, or under circumstances in which things go well. Rather, we should understand these as examples of instances in which things malfunction or go wrong, blocking the transparency that usually exists.

To avoid the dilemma that I outlined above, I propose that democratic procedures' transparency might be similarly blocked. For instance, a decision made by the electorate might contradict another decision made by the same electorate. Then we might argue that an agent may resist treating the procedure as binding in this particular instance, despite acknowledging its transparency under normal conditions. And then we might proceed to identify other formal conditions that would justify an agent judging that transparency had been blocked. For instance, perhaps a decision that distributed civil rights unequally among different races would fail to be transparent. Then an individual might have a formal standard with which to judge whether a democratically made decision actually bound her, normatively speaking—a standard other than one that simply permits her to disobey *whenever* she disagrees.

Now, this is by no means a complete account of how civil disobedience could be justified. And if the conditions that block transparency are going to be anything other than *ad hoc*, then much more would need to be said about why, for instance, distributing civil rights unequally would block transparency. But my view, by explaining how the decisions of the democratic body exert normative force on us, does provide a background against which to ask the question about when such normative force is lacking. And most importantly for my purposes here, there is no reason to assume that an account of how transparency is blocked cannot be given, and so there is no reason to assume that my view cannot escape this putative dilemma.

## 9. Conclusion

My conclusion, then, is that democratic institutions have authority. So when the society decides, democratically, to pursue some end or project, the individual ought to prioritize it. This is not because the balance of reasons favors obedience. It is because whether or not the balance of reasons favors the course of action undertaken by the entire society, the decision of the democratic body is transparent to what the individual does. As such, the normative mistake that a disobedient citizen makes does not consist in misjudging her reasons. Like the akratic agent, her mistake consists in a failure to respond to the normative demands of transparency.

## Chapter 6: Liberalism, Republicanism, and Choice

### 1. Introduction

I have, throughout this dissertation, argued for an expansive role for the state. Specifically, I have argued that one of the state's proper functions is to facilitate collective projects that are good or worthwhile generally, and that members ought to cooperate with the projects that a democratically constituted commonwealth chooses. This is a view of the state and of members' relationship to it that is opposed by most liberals and republicans (and also libertarians).<sup>1</sup> While they differ on many of the details, what unifies philosophers from these traditions is their belief that the individual has a *prima facie* claim to be left to pursue her own individual ends, and to not be enlisted in the pursuit of others' ends or projects—a claim that acts as a general constraint on what the state may do. The coercive powers of the state, they think, are not to be used to facilitate worthwhile collective pursuits generally. The proper role of state institutions involves maintaining independence between members, enabling the individual to make her own choices and pursue her own individual ends. Coercive acts on the part of the state are justified only when they protect the individual's ability to pursue her ends freely, or when they positively facilitate an environment in which she can do so.

Nowhere have I denied that an individual's ability to pursue her own ends should factor into the people's deliberations during the policy-making process. It is, after all, important that

---

<sup>1</sup> For a republican version of this view, see Philip Pettit, *Republicanism: A Theory of Freedom and Government* (Oxford: Oxford University Press, 1999). Kant held that the purpose of the state is to realize a condition of equal freedom, meaning roughly a condition in which independent persons are free to pursue their own ends. His argument is spelled out in the *Metaphysics of Morals*, and is defended at length in Arthur Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy*, (Cambridge, Mass.: Harvard University Press, 2009). Contemporary political liberalism is defined by its commitment to letting each individual pursue her own conception of the good. (See John Rawls, *A Theory of Justice, Revised Edition* (Cambridge, Mass.: Harvard University Press, 1999).) Libertarians begin with an assumption that individuals have a right to determine what they will do with their lives, and that the role of the state is simply to protect that right. (See Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).) There are also hybrid views that combine, for instance, elements of republicanism, liberalism, and libertarianism. (See Friedrich Hayek, *The Constitution of Liberty* (Chicago: University of Chicago Press, 1960).)



an individual have a range of options to choose among as she decides how to live her life. Nevertheless, I have obviously advocated a more extensive role for the state than liberal and republican thinkers tend to allow, and so the liberal and republican views can be aptly described as rivals to the view I have argued for here. While the liberal-republican view of the state's role is probably the dominant view in contemporary political philosophy, my hope is the last five chapters go some of the way toward demonstrating that there exists a plausible alternative. Once we accept that a commonwealth can act, that it is under normative pressure to act in ways that are worthwhile, that it has a good in which all of its members share, and that it can make decisions about what to do that have the force of authority, it should begin to appear that we can no longer take it for granted that the liberal-republican view is true.

I will not, here, attempt to offer a comprehensive critique of liberalism or republicanism. First, that would be an entire project in its own right. But, second, I actually share with these traditions a fundamental concern for self-rule, for respect for the individual person, and for protecting an individual's ability to exercise her rational agency. So I would not want to repudiate these views in their entirety. Instead, I will simply raise some doubts about the liberal-republican view about what the ideal commonwealth should look like, and suggest that some considerations that are often used to buttress the liberal-republican view can be employed to support the vision of the commonwealth that I argued for over the first five chapters.

Rather than arguing against a generic version of the liberal-republican view, I will proceed by examining two authors—Arthur Ripstein and John Rawls. My hope is that by showing how their views fail to pose a threat to my position—and perhaps how parts of their view align with mine—I will give the reader some idea of how my view could be defended from other liberal or republican accounts.

## 2. Ripstein's Kantian View

Unlike some liberals, republicans, and libertarians who simply begin with a presumption in favor of allowing rational adults to decide for themselves what to do, Arthur Ripstein is more interested in exploring the justification for doing so. For this reason, he will make a good proxy for others in the liberal and republican traditions. In this section, I will summarize the set-up work that Ripstein does to motivate his Kantian view of the proper role of state institutions.<sup>2</sup> I will then proceed to argue that his view does not, in the end, force him to endorse the conclusion that the commonwealth is forbidden from using the powers of the state to pursue worthwhile collective ends of its own.

Ripstein begins by saying that, "Your sovereignty, which Kant also characterizes as your quality of being 'your own master (*sui juris*),' has as its starting point your right to your own person, which Kant characterizes as innate."<sup>3</sup> So you have an innate right to your own *person*. A person, according to a fundamental doctrine in Roman Law, "is a being capable of setting its own purposes, [as opposed to] something that can be used in the pursuit of whatever purposes the person who has it might have."<sup>4</sup> As a being with a right to your own person, then, it is *you* who have the ultimate claim to set purposes for yourself. Ripstein describes the upshot, writing that, "The right to be your own master is neither a right to have things go well for you nor a right to have a wide range of options. Instead it is explicitly contrastive and interpersonal: to be your own master is to have no *other* master."<sup>5</sup>

---

<sup>2</sup> I am going to attribute the view to Ripstein, despite the fact that he takes himself to be merely interpreting Kant.

<sup>3</sup> Arthur Ripstein, *Force and Freedom*, p. 35.

<sup>4</sup> *Ibid.*, p. 36.

<sup>5</sup> *Ibid.* Here Ripstein is adopting language that is explicitly republican. For an account of how republicans have traditionally used the notion of being one's own master in order to characterize freedom, see Pettit's *Republicanism*.

Ripstein says that there are several ways that one can come to be dominated by another—come, that is, to serve as someone else’s master. The first is straightforward: you can dominate me by using physical force to coerce me to act, or by dragging my body against my will from one place to another.<sup>6</sup> Another way to dominate me is to place yourself in a position to issue threats against me. You dominate me if, for instance, you kidnap my son and are thus in a position to harm him if I refuse to obey you. There are, however, more subtle ways for one individual to dominate another. For instance, if the economy is constituted in such a way that I depend upon you for charity, then I am in a position to withhold things that you need to live if you refuse to obey me.<sup>7</sup> This, too, is a form of domination.

These examples are all offered by Ripstein, but I believe that there are additional ways in which persons might relate to one another that could also raise worries about domination. For instance, if one person or multiple persons can, through means that would be otherwise unobjectionable, eliminate some or many of the options that an individual has, then they have significant control over what that individual does.<sup>8</sup> To give an example, imagine a media mogul who purchases all competing newspapers and closes them down so that anyone who wishes to pursue a career in journalism will have to do so on terms that the mogul sets. A prospective journalist must answer to the mogul, and this should raise some concerns about domination.<sup>9</sup>

---

<sup>6</sup> Ibid., p. 54.

<sup>7</sup> Ibid., pp. 54-55, pp. 272-286.

<sup>8</sup> Nozick offers an example in which there exists a set of people, each of whom wishes to marry someone within the set. In such a circumstance, it would be possible for all but two members to choose a partner, at which point the two remaining members would have no choice of partner besides the other remaining person. (See Nozick’s *Anarchy, State, and Utopia*, p. 263.) Nozick’s point is that when the two remaining members wed, they do so voluntarily, and are not wronged in having their choice situation constrained in this way. Intuitively, this seems right. But, importantly, the point that I am making is not that having the ability to eliminate other people’s choices is *necessarily* impermissible. Perhaps in the case of choosing romantic partners it is not. Rather, the point is that if we are concerned generally with the ability of individuals to make choices for themselves, then one’s having the ability to control another person’s choice situation in this way raises worries, and at the very least some kind of justification is needed in order to explain why it is permissible for a person to have this sort of control over another.

<sup>9</sup> Ripstein also makes the point that an individual does not have a right to a broad range of options. He does think, however, it is problematic if by limiting my options, you have the ability to prevent me from

Or, a person or persons might exercise control over an individual not by limiting her options *per se*, but by making many of her options so expensive or undesirable that there exists only one *rational* option. So if one has the ability to render various transportation, education, and career options so costly that making use of them would be irrational, then one has the ability to exercise significant control over how others transport themselves, which institutions they study at, and which careers they pursue. This also, it seems to me, raises concerns about domination. Finally, if the members of a society arrange things so that a person cannot participate in the economy without making a contribution to some collective end or project—maybe because participants are taxed, and the taxes are allocated toward some end or to sustaining some group activity—then the society has a way of involving the individual in its collective projects, whether or not the individual chose the activity for herself. Here too, we have some individuals exercising control over others in a way that could raise concerns about domination.

So there are many types of control that individuals can exercise over other individuals, or that groups of individuals could exercise over individuals, that raise worries about the possibility of domination. However, Ripstein makes clear that the mere fact that one is coerced, or that one might be coerced, does not necessarily imply that one is dominated. He explains:

Coercion is objectionable where it is a hindrance to a person's right to freedom, but legitimate when it takes the form of hindering a hindrance to freedom. To stop you from interfering with another person upholds the other's freedom. Using force to get the victim out of the kidnapper's clutches involves coercion against the kidnapper, because it touches or threatens to touch him in order to advance a purpose, the freeing of a

---

interacting with others and entering into voluntary relationships with them (*Force and Freedom*, pp. 245-246.) But I think the problem runs a little bit deeper than this. If others can narrow my options so that I only have a single choice, then others have control over my choice. The worry is not about having a broad range of options *per se*, but about others having control over what I do.

victim, to which he has not agreed. The use of force is rightful because an incident of the victim's antecedent right to be free.<sup>10</sup>

Coercion is rightful, Ripstein says, when it is necessary to restore "equal freedom,"<sup>11</sup> which is a condition in which no person is in a position to exercise control over others that the others are not able to exercise over her. Ripstein, following Kant, thus believes that equal freedom serves as a standard by which to evaluate whether various forms of coercion are legitimate.

The picture that emerges, as I understand it, is this. As a distinct individual, each has an innate right to her own person. A person is essentially an end-setter, and so the innate right of humanity is a claim to set one's own ends and purposes. The state uses its coercive power legitimately if and only if it does so in the service of realizing equal freedom—realizing, that is, a public sphere within which everyone is free to set and pursue her own ends and purposes, while at the same time respecting the right of others to do the same.

The state's proper function thus involves both preventing people from physically coercing one another in ways that are wrongful, and creating conditions in which people are not in a position to coerce one another in ways that undermine equal freedom. One thing that is clear, however, is that the state is not in the business of selecting and facilitating collective pursuits simply because they are worthwhile. To coerce citizens through force, or to tax them, or to coerce them by limiting all their other options for the purpose of extracting a contribution to some collective pursuit, would be to coerce persons for reasons other than the realization of equal freedom. This is exactly the kind of coercion that Ripstein claims is wrongful.

So, now that I have laid out the basics of Ripstein's account, I want to argue that what he says does not actually warrant the conclusion that the state ought not work to facilitate collective pursuits generally. Obviously the innate right of humanity plays a fundamental role in Ripstein's account. Unfortunately, Ripstein's account contains scant discussion about what reason there is

---

<sup>10</sup> Ibid., p. 55.

<sup>11</sup> Ibid.

to think that individual persons possess this right, and so we are left to speculate for ourselves about this. So here I will consider two possible reasons for positing the innate right of humanity. Then I will argue that if the innate right of humanity is grounded in either of these ways, then there is reason to suspect that the state's role is not limited in the way Ripstein thinks it is.

The first possibility is that it is simply a "brute fact" that beings or entities with the capacity for choice or purposive activity have an innate right to exercise it. Ripstein does, at some points, seem to be relying upon an idea like this, as he denies that the right is derived in the way that moral norms are in Kant's ethical works,<sup>12</sup> and he often talks as if it is just a ground-level fact in political philosophy that end-setters possess this right. But if Ripstein is, indeed, relying on this brute fact, then we ought to consider an entire people or society to be a bearer of this right as well. This is because, as I argued in Chapters 1 and 2, an entire society has the capacity to engage in purposive activity. And, as I argued in Chapter 5, a people can, through democratic procedures, choose activities for themselves. Of course, there are important differences between individual agents and group agents. A group is not a biological organism or a member of the species *homo sapiens*, for instance. But certainly Ripstein, as a Kantian, does not want to ground the innate right of humanity in a person's biology. So if we rely on the brute fact that beings or entities with the capacity for purposive activity and choice are bearers of the innate right of humanity, then I do not see a way around the conclusion that a people is a bearer of that right as well.

A second possibility for grounding the innate right of humanity would be to appeal to the important role that action plays in an agent's life. The lives of rational agents are partially constituted by action, and so it is very important to the individual *which* actions she engages in. Or perhaps we might say that as beings with the capacity for purposive activity, the *use and exercise of the capacity* itself is very important to us. For either of these reasons, a right that protects a person's ability to choose her own actions might be thought to be important. I actually

---

<sup>12</sup> Ibid., pp. 11-13.

doubt that Ripstein would agree that this consideration plays a role in grounding the innate right of humanity. But others might think this, and since my interest here is not just Ripstein *per se*, but the family of views he represents, I will consider this possibility.

To make the argument, I am going to take a bit of a detour and consider the individual agent, taken in isolation. It is uncontroversial to assert that an individual exerts significant control over her own future. This means that it is often within an individual's power to manipulate and plot against herself. The individual can make choices now with an eye to limiting the options she will have in the future. She can make choices now with an eye to limiting the *rational* options she will eventually have. She can get involved in a project, activity, or relationship that she knows she will eventually be unhappy about, but also knows it will be difficult to extricate herself from. In these respects, the individual is in the same position to exercise control over her future as some individuals are to exercise control over distinct individuals. In Chapter 3, I argued that an individual has a moral duty to respect her own autonomy. Because she has such a duty, and because she has so much control over her future, the individual faces a predicament that parallels the commonwealth's predicament—she faces the question of how (or whether) she can act without violating the requirement to respect her autonomy and to respect the ends she will have in the future.

Now, it isn't tempting at all to suggest that an individual should try, as much as possible, to leave it open to her future self to decide what to do. This is because when we say that it matters whether an individual is able to exercise her capacity for choice, we are generally not concerned with her ability to will isolated, momentary movements, that have little to do with what precedes them or what follows. That is to say, it isn't as if the value we place on autonomy or on setting our own ends stems merely from the importance of being able to wiggle our toes or twiddle our fingers. Defenders of freedom and autonomy are usually most anxious to point out that it matters whether an individual can choose her career, her avocations, what kinds of relationships to carry on with, which church to participate in, which intellectual projects to

pursue, and so on. But all of these choices seem to involve long-term commitment—and crucially, they involve integrating many portions of an individual's life, rather than keeping them independent, or attempting to leave decisions open until later. So when an individual chooses a religion, or a life partner, or to have children, or to pursue a particular career, she closes off options for the future, makes irrational many options that would have otherwise been rational, and commits herself to projects from which it is difficult to extricate herself. But were we to attempt to leave open as much as possible what we will do in the future, we would certainly undermine our ability as individuals to carry on with much meaningful activity at all. We would, that is, destroy much of what was valuable about engaging in activity to begin with. So whatever solution we settle upon when trying to determine how an individual can avoid wronging herself when she exercises control over her future self, it must be a solution that allows an individual to integrate her life activity in order to carry out worthwhile projects or actions.

With that said, let us return to Ripstein. We were supposing that, in order to ground the innate right of humanity, we might appeal to the fact that it matters which actions an individual engages in or that it matters that she be able to exercise her capacity for choice. However, as far as I can tell, there does not seem to be good reason to deny that it matters which *group actions* an individual engages in. Insofar as an individual cares about what she is doing, shouldn't it also matter whether she is engaged in a collective act of war, a collective effort to sustain just economic practices, or a collective effort to realize the common good? But if, in order to avoid wronging individual members, we attempt to realize a society in which each individual cannot be coerced or drawn into the society's collective activity, then we would undermine our ability to carry on with these important collective actions. That is to say, just as the importance of being able to engage in worthwhile action recommends against an individual's trying to leave it open to her future self to decide what to do, it recommends against trying to maintain independence between members of a society. Given that individuals often do not act as they ought because



they are tempted to free ride or because they are selfish, the possibility of doing anything at all together as a society hinges on the ability of the whole to exercise some form of coercion. Thus, whatever solution we settle upon when trying to determine how the commonwealth can avoid wronging its members, it should be a solution that allows members to use some form of coercion in order to integrate their collective activity. By no means do I think that this is a decisive case against Ripstein and the Kantian view that he lays out. My point is that once we take seriously the view that societies are agents, capable of pursuing ends and projects, then complications are raised about whether the state's role is as limited as Ripstein has suggested it is.

### 3. Rawls's Political Liberalism

Now I will turn to Rawls's political liberalism. As I understand him, Rawls conceives of a liberal society as a scheme of social cooperation within which each is free to pursue her own conception of the good life. Collective decisions are necessary in order to define the terms under which everyone will cooperate—to determine, for instance, how burdens and benefits will be distributed. But collective decisions are *not* made with an eye to deciding what it would be good generally for the society to pursue together, as a whole. The notion of *our* pursuing the good together, or *our* living out a collective life, plays no role on Rawls's account. Famously, Rawls holds that the justification for state action may not be based on any controversial view about the good or about what good action would consist in. So when deciding which tax code to institute, or which projects to allocate funds to, citizens do not make the decision on the basis of what would be best overall, or by considering what the balance of all reasons favors. Instead, they limit themselves to public reasons, which means that “political values alone” settle questions pertaining to state coercion.<sup>13</sup> Rawls means a number of things by this. But one of them is that

---

<sup>13</sup> John Rawls, *Political Liberalism*, expanded paperback edition (New York: Columbia University Press, 1993), pp. 214-215.

within a liberal society, coercion is used to ensure that each person is free to pursue her own conception of the good life, and not to enlist her in worthwhile pursuits.

Most of Rawls's effort is spent spelling out what a politically liberal society would look like, rather than justifying political liberalism itself. But he offers clues throughout as to why he favors liberalism. The importance of stability in a pluralistic society is obviously one reason to endorse political liberalism. Another has to do with the recognition that disagreement occurs between reasonable people due to the, "many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life."<sup>14</sup> However, it seems clear to me that one of the most important grounds for political liberalism is that liberalism shows respect for individual persons as bearers of the power to select and pursue a conception of the good life. Martha Nussbaum, in an explication and defense of Rawls's liberalism, writes:

[Citizens in a liberal society] respect their fellow citizens, and respect them as equals. Their reasonableness is an ethical reasonableness: respecting their fellow citizens, they want to give them plenty of space to search in their own way, even though they may believe that the conclusions most people come to are wrong. Respect is for persons, not directly for the doctrines they hold, and yet respect for persons leads to the conclusion that they have liberty to pursue commitments that lie at the core of their identity, provided that they do not violate the rights of others. . .<sup>15</sup>

---

<sup>14</sup> Ibid., p. 56.

<sup>15</sup> Martha Nussbaum, "Perfectionist Liberalism and Political Liberalism," *Philosophy and Public Affairs* 39 (2011), pp. 3-45 (quotation from p. 17, footnote omitted).

So political liberalism leaves a person to decide for herself what to do not out of respect for the choices that she makes, but instead out of respect for her as a person—a person who bears the capacity to select ends for herself.

It is for this reason, I believe, that political liberalism treats citizens as, “self-authenticating sources of valid claims.”<sup>16</sup> So when, for instance, a Catholic chooses to pursue her own salvation in a particular way, in so doing places a claim on her society’s institutions not to interfere. When the state respects her freedom in this way, it is not endorsing any particular view about salvation. It is, instead, a sign of respect for the person herself. And if the Catholic wishes to coerce another person—a non-believer, let us say—into pursuing salvation with her, the fact that the liberal social institutions forbid this is not taken as an indication that the society is taking sides about how it is best to live, but is rather taken as a sign of respect for the non-believer as a person, regardless of the truth or falsity of her beliefs.

Some have doubted whether it is possible to respect individual persons in the way that Rawls envisions, claiming that any coercive system of laws will imply that some conceptions of the good are true and others are false.<sup>17</sup> But rather than raise this worry, I want simply to question whether Rawls’s version of political liberalism is the best way to respect individuals. Throughout this dissertation, I have argued that an individual’s action, reasons, and good are deeply entangled with her society’s action, reasons, and good. What an individual does is partly constituted by what her society does. What she is under normative pressure to do is partially determined by what her society is under normative pressure to do. Part of what is good for her involves participation in a commonwealth’s collective endeavors. The fact that she helps to constitute a society with particular features not only puts her under the authority of the General Will, but gives her authority as a member of the body that decides what to do using the General Will. If all of this is right, then at the very least it should not be taken as a foregone conclusion

---

<sup>16</sup> Rawls, *Political Liberalism*, p. 32.

<sup>17</sup> Michael Sandel, “Review of Political Liberalism,” *The Harvard Law Review* 107 (1994), pp. 1765-1794.

that the proper way to respect persons is to merely let them live out their own individual conceptions of the good life. It should not be taken as a foregone conclusion that the proper way to respect individuals is to, “give them plenty of space to search in their own way,” as Nussbaum claims. Given the arguments I have made throughout the dissertation, we might think that the best way to give respect to each individual is to give her an equal role in the process that determines what a people will do together, and how they will go about doing it. We might think it more respectful not to banish conceptions of the good from public deliberation, but to give each an opportunity to propose and advocate for her views about how best to realize the good. After all, in Chapter 2 I argued that individuals *just are* members of the society, and in Chapter 5 I argued that as such their judgments about what to do have the force of authority (though no more or less authority than each other person’s judgment). So we might think that respect requires not that we give each person space to search in their own way, but that we give each a voice when determining what “we” will do together. Respect might require treating the democratic body itself as a self-authenticating source of valid claims. The notion of respect, here, is vague enough that I am not entirely sure how to adjudicate a dispute about how best to be respectful. My suggestion is merely that if one accepts the arguments I made earlier in the dissertation, then it becomes less obvious that the correct way to respect individuals is the way Rawls proposes.

To clarify, and to avoid a possible misunderstanding, I am not claiming here that Rawls has made a metaphysical mistake by failing to take into account the fact that the individual is essentially a member of a commonwealth, that the General Will has authority over her, or that her good is intertwined with others’. After all, Rawls does not think that political liberalism takes a stand on the truth of any of these controversial metaphysical doctrines. His point is that by not taking a stand on the truth of these controversial metaphysical doctrines, we give respect to each member. What I am suggesting is simply that given what I have said throughout the dissertation, we should not be so confident that this is the proper way to give respect to each.

Finally, I should say again that none of this is meant to recommend that the state should be involved in every aspect of our lives or that individuals should be left no space at all within to pursue their own conception of the good. Everything that I say here is consistent with the claim that people should have wide leeway to worship as they wish, pursue careers of their choosing, marry who they are attracted to, and live where they choose. My argument simply implies that providing an environment in which these things are possible is not the only thing that a state that respects the individual ought to do. It ought also pursue collective projects that it deems worthwhile.

#### 4. Conclusion

I said in the dissertation's introduction that I would advocate a Rousseauvian vision of an ideal society: one in which the people constitute a group agent, unified under a collective will, willing action that constitutes the common good. Most of the dissertation was a positive argument for that vision. This chapter, on the other hand, has been mostly defensive. What I have tried to show is that if one accepts the arguments I have made in previous chapters, then it should complicate one's views about which rights we think a person should have and about the proper way to respect an individual—even if one is partial to a liberal or republican view.

Rousseau imagined that a people would unify and act as a people, and that when they did, they would share in the good of the whole. Like Rousseau's view, mine is a bit difficult to characterize in terms of its relationship to liberalism and republicanism. My commitment to self-rule and to respect for individuals echoes both of these traditions. On the other hand, my views about there being a General Will and a common good that everyone shares in sound more communitarian or collectivist. So I expect that the view will inspire many of the same worries that Rousseau has always inspired—worries about totalitarianism and about the individual being swallowed up by the collective. But for some of the same reasons that thinkers have worried about Rousseau's view, many have also found him inspiring. Despite his concern for the

individual, he nevertheless recognized the importance of living out a life together as a people, and argued that we cannot understand the individual without understanding her as a member of a community, society, or collective. His view, thus, has the virtue of integrating commitments and values from several traditions. My hope has been to argue for a view that has this virtue as well.